UNIVERSITY OF CALIFORNIA

Los Angeles

Measurement of Electroweak Single Top Quark Production in Proton-Antiproton Collisions at 1.96 TeV

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Physics

by

Peter Joseph Dong

2008

The dissertation of Peter Joseph Dong is approved.

_____
Zvi Bern

_____
Lincoln Chayes

_____
Jay Hauser

_____
Rainer Wallny, Committee Chair

University of California, Los Angeles

2008

For the invisible things of Him from the creation of the world are clearly seen,

being understood by the things that are made,

even His eternal power and Godhead.

—Romans 1:20a

Table of Contents

List of Figures

List of Tables

Acknowledgments

In 1985, Timothy Ferris made a documentary called *The Creation of the Universe* which was broadcast on PBS. My parents recorded it on our Betavision and I watched the tape at least a hundred times and decided that I had to be a physicist. So thank you, Mr. Ferris, wherever you are, for starting me on the course that ended up here.

I am very grateful to the legions of excellent physicists and engineers in Fermilab's Accelerator Division who keep the accelerators running smoothly, and to the additional legions of excellent physicists at CDF who are responsible for the design, construction, and operation of the detector, as well as the offline processing, reconstruction algorithms, particle identification, scale factors, and all the other essential studies without which this analysis would not be possible.

I am particularly indebted to Bernd Stelzer, who first applied the matrix element method to the single-top search, with whom I shared an office as well as all the frustrations and triumphs (but more of the former) of this analysis. His friendship and support through many late nights at the lab made this dissertation possible. I also would like to thank Tom Junk for many extremely fruitful discussions about statistics, as well as for writing and explaining the software that performs the final calculations; Charles Plager, for four years of banter, arguments, and Snapple Real Facts [1] in the CDF lunch room; Ricardo Eusebi, without whom the diamond radiation monitoring system would never have been completed,

which would mean this analysis would never have been completed; and Florencia Canelli, who wrote the first matrix element thesis, for all her advice and support. I would also like to thank the people who helped me get my bearings when I first started at CDF, especially Rick Tesarek, George Wyatt, Mike Lindgren, and JJ Schmidt.

Of course I need to thank my adviser, Rainer Wallny, who kept supporting me and believing that I could do great things, even when there was evidence to the contrary—such as the fact that I was always falling asleep in the office. Under his guidance, I learned how to design and install a hardware system (though I was never very good at it), how to program, how to prioritize and multitask, how to understand the factorization scale, how to deal with a large collaboration, and how to find good restaurants near the lab. I particularly appreciate the chance to drive his very nice car across the country.

Most of all, I would like to thank my family, who always supported my interest in physics, especially my father, who was unable to complete his own Ph.D. in physics at UCLA. In some sense, this dissertation is his accomplishment as well.

VITA

December 10, 1979	Born, Melrose Park, Illinois
2002	B.A., Physics and Music Harvard University Cambridge, Massachusetts
2002–2003	Teaching Assistant Department of Physics and Astronomy University of California, Los Angeles
2003–2004	Graduate Student GK-12 Fellow National Science Foundation University of California, Los Angeles

PRESENTATIONS

P. Dong, "Search for Electroweak Single Top Production Using a Matrix Element Method at CDF," presented at the meeting of the American Physical Society (April 2007), Jacksonville, Florida.

P. Dong, "Beam Condition Monitoring with Diamonds at CDF," presented at the IEEE Real-Time Conference (April 2007), Batavia, Illinois.

P. Dong, R. Eusebi, C. Schrupp, A. Sfyrla, R. Tesarek, R. Wallny, "Beam Condition Monitoring with Diamonds at CDF," IEEE Trans. of Nucl. Sci. **55** (2008) 328–332.

P. Dong, "Beam Condition Monitoring with Diamonds at CDF," presented at the Fermilab Users' Meeting (June 2007), Batavia, Illinois.

P. Dong, "Single Top Quark Physics," presented at Les Rencontres de Physique de la Vallée d'Aoste (February 2008), La Thuile, Italy.

P. Dong, "Search for Electroweak Single Top Production Using a Matrix Element Method at CDF," presented at the meeting of the American Physical Society (April 2008), St. Louis, Missouri.

ABSTRACT OF THE DISSERTATION

Measurement of Electroweak Single Top Quark Production

in Proton-Antiproton Collisions at 1.96 TeV

by

Peter Joseph Dong

Doctor of Philosophy in Physics

University of California, Los Angeles, 2008

Professor Rainer Wallny, Chair

The top quark is an extremely massive fundamental particle that is predominantly produced in pairs at particle collider experiments. The Standard Model of particle physics predicts that top quarks can also be produced singly by the electroweak force; however, this process is more difficult to detect because it occurs at a smaller rate and is more difficult to distinguish from background processes. The cross section of this process is related to the Cabbibo-Kobayashi-Maskawa matrix element $|V_{tb}|$, and measurement of the single top quark

production cross section is currently the only method to directly measure this quantity without assuming the number of generations of fermions.

This thesis describes a measurement of the cross section of electroweak single top quark production in proton-antiproton collisions at a center-of-mass energy of 1.96 TeV. This analysis uses 2.2 fb^{-1} of integrated luminosity recorded by the Collider Detector at Fermilab. The search is performed using a matrix element method which calculates the differential cross section for each event for several signal and background hypotheses. These numbers are combined into a single discriminant and used to construct templates from Monte Carlo simulation. A maximum likelihood fit to the data distribution gives a measurement of the cross section. This analysis measures a value of $2.2^{+0.8}_{-0.7}$ pb, which corresponds to a value of $|V_{tb}| = 0.88^{+0.16}_{-0.14}(\text{experimental}) \pm 0.07(\text{theoretical})$. The probability that this result originates from a background fluctuation in the absence of single top production (*p*-value) is 0.0003, which is equivalent to 3.4 standard deviations in Gaussian statistics. The expected (median) *p*-value as estimated from pseudo-experiments for this analysis is 0.000003, which corresponds to 4.5 standard deviations in Gaussian statistics.

Chapter 1

Introduction

The Standard Model of particle physics is a very successful theory of the interaction of fundamental particles [2]. The current formulation of the Standard Model includes three sets, or generations, of fundamental particles called fermions. There is, however, no intrinsic reason that there must be only three generations, though a fourth has not yet been observed. Only one Standard-Model process, the charged weak interaction, is capable of transforming a single particle from one generation to another. The probability of this occurring is parameterized by a matrix called the Cabbibo-Kobayashi-Maskawa (CKM) matrix.

Most elements of the CKM matrix have been precisely measured. However, the element $|V_{tb}|$, which represents the probability that a top quark will change to a bottom quark through electroweak interactions, has only recently been measured directly. Limits on other measurements indicate that its value must be very close to one if there are only three generations of fermions. Because the total probability must be unity, a direct measurement of $|V_{tb}|$ that is significantly less than one would indicate the existence of a fourth generation [3].

Measuring $|V_{tb}|$ requires an investigation of the top quark, the most massive known fundamental particle. The only facility currently capable of generating the energy needed

to produce top quarks is the Tevatron particle accelerator at the Fermi National Accelerator Laboratory in Batavia, Illinois. However, top quarks are predominantly produced in pairs by the strong force. Measuring the electroweak coupling $|V_{tb}|$ requires examining a less common process, *single top production*, in which only one top quark is produced. The cross section of this interaction is proportional to $|V_{tb}|^2$; thus, a measurement of the cross section of electroweak single top production allows a measurement of $|V_{tb}|$.

However, studying single top quark production involves many experimental challenges. Single top production occurs at less than one half the rate of top pair production, which is already a rare process. At the same time, the background processes which look similar to single top production occur more than ten times as frequently. Simple experimental techniques are not sufficiently sensitive to measure a single top signal in this sample; the more advanced matrix element method is used to separate the single top signal from the background and extract the cross section.

This thesis presents a measurement of single top quark production in proton-antiproton collisions at the Tevatron with an integrated luminosity of 2.2 fb^{-1}. This chapter presents a short summary of the Standard Model and its basic framework. Chapter 2 describes the experimental apparatus of the Tevatron and the Collider Detector at Fermilab. Chapter 3 describes the generation of simulated events for use in this analysis. In Chapter 4, the reconstruction of detected particles is discussed. Chapter 5 presents the requirements used to select candidate events for this analysis. Chapter 6 describes the modeling of

background processes, while Chapter 7 discusses the estimate of the rate of each process. The matrix element method is introduced in Chapter 8, and Chapter 9 applies it to the data to measure the single top production cross section and $|V_{tb}|$. The last chapter discusses potential future improvements of the analysis and presents an application of this analysis technique to search for the Higgs boson, which is the last remaining particle of the Standard Model not yet detected.

1.1 The Standard Model

The best current knowledge of particle physics is given by a theoretical framework called the *Standard Model* [4], a quantum field theory that explains the electromagnetic and nuclear interactions between particles as resulting from the introduction of local symmetries into the Lagrangian. The Standard Model requires a Lagrangian that is invariant under local $SU(3) \times SU(2) \times U(1)$ transformations. Each of these groups governs the interactions involving a particular type of particle called a vector *boson*: $U(1)$ symmetry describes the *electromagnetic force*, mediated by the massless, electrically neutral *photon*; $SU(2)$ symmetry describes the *weak nuclear force*, mediated by the very massive charged W bosons and the neutral Z boson; $SU(3)$ symmetry describes the *strong nuclear force*, mediated by the massless and electrically neutral *gluon* [5]. Each type of boson only interacts with particles that manifest a particular conserved quantity: photons interact with particles that have electrical *charge*, weak bosons interact with particles that have left-handed chirality, and gluons interact with particles that have a quantum number called *color*.

The remaining fundamental force, *gravity*, has not yet been successfully introduced into the Standard Model; however, since its strength is nearly thirty orders of magnitude less than that of the weak force [6], it can safely be ignored in particle interactions. These symmetries were not derived but chosen, over a long period of theoretical trial and error, to match experimental data. The Standard Model features a host of particles, whose names, charges, and masses are shown in Table 1 [7].

Standard Model particles are of two main types: *fermions*, which have half-integer spin (½ for all observed fermions); and *bosons*, which have integer spin (1 for all observed bosons). There are two known types of fermions: *quarks*, which have fractional charge and interact by the strong force as well as the electromagnetic and weak forces; and *leptons*, which have integral charge and interact only by electromagnetic and weak forces. These fermions are divided into three families, each containing two quarks—one with electric charge $+\frac{2}{3}$ and one with charge $-\frac{1}{3}$—and two leptons, one with a charge of -1 and one *neutrino*, which has no electric charge and almost no mass.

Each particle in the Standard Model has an *antiparticle* which is identical in all regards but with opposite charge. Because of their similarity, in this document particles and antiparticles are usually referred to inclusively: "electron" refers to an electron or a positron, "muon" refers to a muon or an antimuon, and so forth.

Because they have no color, charged leptons interact by the electromagnetic and weak forces but not the strong force. Neutrinos have no charge and thus only interact by the

	Particle (abbreviation)	Charge	Mass (GeV/c^2)
Leptons	Electron (e)	-1	$5.10998918 \pm 0.00000044 \times 10^{-4}$
	Electron neutrino (ν_e)	0	$< 1.9 \times 10^{-10}$
	Muon (μ)	-1	$0.1056583692 \pm 0.0000000094$
	Muon neutrino (ν_μ)	0	$< 1.9 \times 10^{-10}$
	Tau (τ)	-1	1.7769 ± 0.00020
	Tau neutrino (ν_τ)	0	$< 1.9 \times 10^{-10}$
Quarks	Up (u)	$+\frac{2}{3}$	0.0015 to 0.0030
	Down (d)	$-\frac{1}{3}$	0.003 to 0.007
	Strange (s)	$+\frac{2}{3}$	0.095 ± 0.025
	Charm (c)	$-\frac{1}{3}$	1.25 ± 0.09
	Bottom (b)	$+\frac{2}{3}$	4.70 ± 0.07
	Top (t)	$-\frac{1}{3}$	172.5 ± 2.7
Gauge bosons	Photon (γ)	0	$< 6 \times 10^{-28}$
	W boson (W)	± 1	80.304 ± 0.029
	Z boson (Z)	0	91.1876 ± 0.0021
	Gluon (g)	0	0
	Higgs boson (H)	0	129^{+74}_{-49}

Table 1. Properties of the known particles in the Standard Model. Each of these particles has an antiparticle with opposite charge but the same mass and spin. Because gluons cannot be observed directly, the theoretical massless value is used. The Higgs boson has not yet been observed, and its mass is predicted from other electroweak quantities.

weak force. Quarks and gluons are the only particles in the Standard Model that interact by the strong force. The strong force differs from the other two forces in that it exhibits *asymptotic freedom*: the strength of its coupling decreases with increasing energy. Thus, if a quark receives a large amount of energy in a particle interaction, as it moves away from other partons, the binding energy between them increases. When the energy is sufficiently large, a new particle-antiparticle pair is created. This process repeats until the quark's energy is expended, resulting in not a single particle but a collection of particles moving in the same direction, called a *jet* [8]. The quarks couple together to form bound states which are color singlets. The only observed bound states, called *hadrons*, contain a quark and an antiquark—called a *meson*—or three quarks or antiquarks—a *baryon*. Because the coupling of the strong force is much larger than the other two forces, the strong force tends to dominate quark interactions.

Charged weak interactions are the only interactions that change quark flavor. Most of the time, quarks couple to the corresponding quark in their family; for example, a top quark will couple to a bottom quark and a charged weak boson. However, there is a small probability that the quark may couple to a quark in a different family. For example, a bottom quark may couple to an up quark. This probability is determined by the *Cabbibo-Kobayashi-Maskawa* (CKM) *mixing matrix*, which gives the amplitudes for coupling between different flavors in charged weak interactions [9]. The CKM matrix is usually written in terms of the coupling of the mass eigenstates d', s', b' to the flavor eigenstates d, s, and b:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

However, the values of the CKM matrix have no theoretical prediction and must be measured experimentally. Measuring these parameters is a fundamental part of understanding the Standard Model.

1.2 Cross section calculations

The probability of a given interaction occurring is related to the *cross section* of the interaction, which is measured in units of *barns* (b), where one barn is 10^{-24} cm^2. Current theory is unable to calculate these cross sections exactly; instead, a perturbative expansion must be made in powers of a coupling constant. These perturbative terms can be conveniently represented by *Feynman diagrams*, which are graphical representations of each term in the expansion. Feynman diagrams consist of lines, representing fields, and vertices, representing the interactions of the fields. These combine according to a simple set of rules, and the lines and vertices of a Feynman diagram can be converted directly into a calculation of the term in the cross section associated with that diagram. The sum of all terms gives the amplitude of the process [10].

For a given set of initial- and final-state particles, the Feynman diagrams with the fewest possible number of vertices represent the *leading-order* term of the perturbative expansion, and often constitute a good approximation of the underlying physics. One

pleasant feature of most leading-order diagrams is that they contain no closed loops; these are referred to as *tree-level* diagrams. *Next-to-leading-order* diagrams have at least one more vertex and represent the next term in the expansion. Calculating these is much more difficult because of the properties of *loop* diagrams. Whenever the topological feature of a loop appears, the calculations require an integral to be performed over the momenta of the particles in the loop, and the integral often diverges. This is not, of course, a problem with reality, but an artifact of perturbation theory: the next-to-leading-order term of a perturbative expansion is a theoretical construct and cannot be measured. These divergences, usually called *ultraviolet divergences* because they occur for very large momentum scales, can be dealt with by a process called *renormalization* [11], in which a *renormalization scale* is introduced to truncate the integral before it diverges. (It can be thought of as the region in which the theory is valid.) This gives a finite result to the calculation that agrees well with experiment for many interactions.

The introduction of a renormalization scale causes the coupling between fields, initially taken as a constant quantity, to change at different energy scales. The *running coupling constant* introduces a problem for practical calculations because the energy at which the coupling constant is calculated must be chosen; it cannot be derived from the interaction itself. This ambiguity causes some uncertainty in the calculation of cross sections or the generation of simulated events.

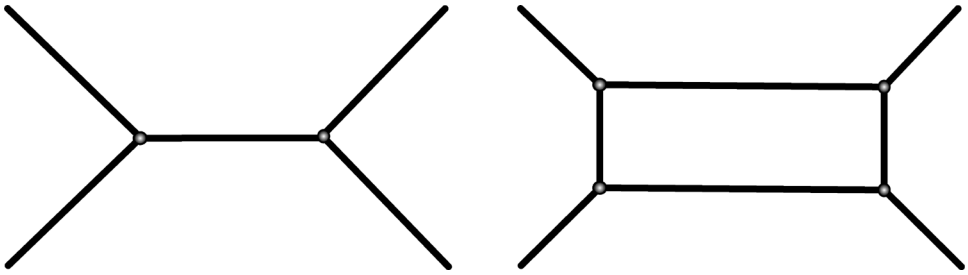

Figure 1. Examples of Feynman diagrams: (left) a tree-level diagram, and (right) a loop diagram. In this work, Feynman diagrams are always drawn with time on the horizontal axis, increasing from left to right.

Another kind of divergence is called an *infrared divergence* and arises for small momentum scales. The source of these divergences is the perturbative nature of quantum field theory calculations. In calculations involving the strong force, at low energies the coupling constant becomes larger than one. In this case, each successive term in the perturbation is larger than the one before it, and perturbation will no longer give a valid answer. Such divergences are dealt with by introduction of a *factorization scale* which truncates the integral before it reaches the nonperturbative region [12]. Fortunately, the energies of particles in this analysis are well above the cutoff for valid perturbation; unfortunately, the choice of renormalization and factorization scale affects the cross section calculation, so it sometimes requires the addition of a systematic uncertainty.

The sum of all possible connected Feynman diagrams gives, in the end, a number related to the amplitude of a given initial state changing to a given final state. It can be imagined as one element in a matrix (usually called the *S*-matrix) which includes the amplitudes of all possible initial and final states, and so is usually called the *matrix element*.

However, it includes only *dynamical* information about the interaction between the particles; it does not include the *kinematic* information that comes from the energy and momentum of the particles. It is convenient to calculate these two terms separately [13].

Because the cross section of a particle with given exact values of momentum and energy is infinitesimally small, it is more sensible to calculate the *differential cross section* of the interaction. This quantity is defined for an infinitesimal slice of the momentum space of all final state particles. For a cross section calculation, it is given by Fermi's Golden Rule: for scattering of two particles with four-momenta q_1 and q_2 into n particles with four-momenta p_i, the differential cross section is given by

$$d\sigma = |M|^2 \frac{\hbar^2 S}{4\sqrt{\left(q_1 \cdot q_2\right)^2 - \left(m_1 m_2 c^2\right)^2}} d\Phi_n$$

where S is a combinatorial factor for identical particles, m_i is the particle mass,

$$d\Phi_n = \delta\left(q_1 + q_2 - \sum_{i=1}^{n} p_i\right) \prod_{i=1}^{n} \frac{c d^3 p_i}{(2\pi)^3 2E_i}$$

is the phase space factor, and M is the matrix element for the interaction [14]. Integrating this expression for all final-state momenta gives the total cross section of the interaction.

1.3 The top quark

The quarks in the first family, the up and down quark, make up protons and neutrons, which together with electrons make up most of the visible matter in the universe. The other

quarks are unstable and decay rapidly upon creation. Because of the nature of the strong force, in which the coupling increases with increased separation, a lone quark can never be observed, and its mass is difficult to determine or even define; the masses of light quarks are usually given as estimates [15]. The masses of the fermions have not been explained by any theory and are currently taken as input parameters in the Standard Model.

The top quark is the only exception to the mass measurement problem. The mass of the top quark is extremely large—at more than 170 GeV/c^2, it is far more massive than any other known elementary particle. Its large mass makes its lifetime extremely short (roughly 10^{-26} s), so short that it decays before it has a chance to form into hadrons. This makes it easier to probe some of its properties, such as its spin, its mass, and its lifetime.

The top quark was first discovered in Run I of the Tevatron at Fermi National Accelerator Laboratory (Fermilab) in proton-antiproton collisions with a center-of-mass energy of 1.8 TeV [16]. Subsequent measurements in Run II of the same facility have improved the understanding of its production, decay, mass, and charge.

1.4 Single top production

Top quarks are typically produced by the strong force: a quark-antiquark pair form a very energetic gluon, which then decays into a top-antitop pair (Figure 2). Since the strong force has a stronger coupling than the other forces, it dominates the production of top quarks. However, top quarks can also be produced through electroweak interactions. Charged weak-

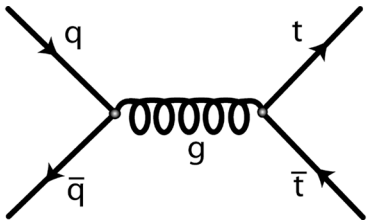

Figure 2. The Feynman diagram for top pair production.

current interactions can mediate the production of a top-antibottom pair, since these interactions change quark flavor.

Leading-order diagrams for single top production are shown in Figure 3. The *s*-channel and *t*-channel diagrams are the dominant production channels at the Tevatron; associated *W* production is not expected to be observable at the Tevatron with the amount of data collected to date.

Measurement of the single top production cross section allows the measurement of several Standard Model parameters. The most important of these is the CKM matrix element $|V_{tb}|$, which is the probability amplitude that a top quark will couple to a bottom quark in a charged weak interaction, and can only be directly measured in single top production. Other measurements of $|V_{tb}|$ have been made by studying the rate of top quark decays [17], but these measurements assume three families, because if a fourth family of heavier quarks existed, the top quark would be unable to decay to these particles. In the case of single top production, however, the cross section is directly proportional to the square of $|V_{tb}|$; measuring the single top production cross section gives a measurement of $|V_{tb}|$ that makes

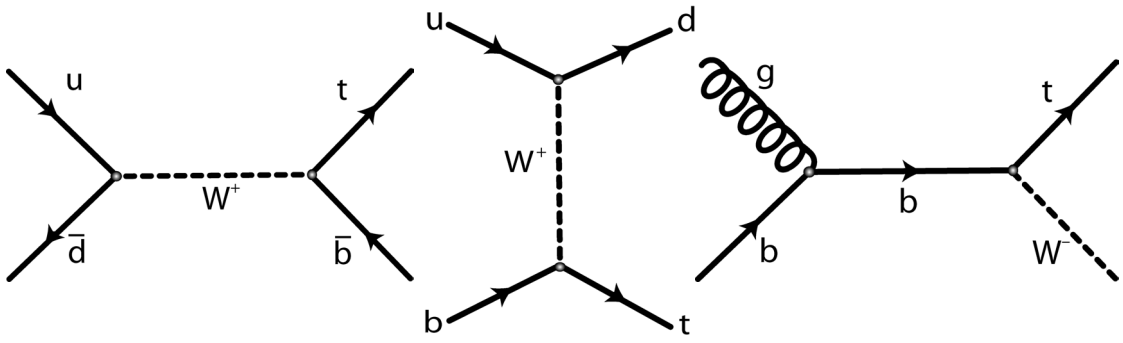

Figure 3. Representative Feynman diagrams for *s*-channel (left), *t*-channel (center), and associated *W* (right) single-top production. Similar diagrams with all particles changed to antiparticles are also possible.

no assumption about the number of fermion families.

In addition, the Standard Model predicts that the top quarks resulting from this interaction will be almost entirely polarized, since the *W* boson only interacts with left-handed particles [18]. This polarization allows a probe of the spin projection of the top quark and the chirality of the *W* boson.

Discovering single top production is also an important milestone in the search for the Higgs boson. The signature of *WH* production, which is the most sensitive mode for a low-mass Higgs boson at the Tevatron [19], has the same final state as single top production if the *W* boson decays leptonically. Searches for the Higgs boson face similar challenges to single top searches, so they will most likely employ similar techniques. A single top production measurement is thus a proving ground for the sophisticated analysis methods needed to observe the Higgs boson.

Any significant deviation from the theoretical value of the cross section of single top production could be an indication of physics not predicted by the Standard Model. For example, a value lower than the predicted value of the cross section might indicate that $|V_{tb}|$ is less than one, which could indicate the presence of a fourth family of fermions. Because studies of branching ratios of Z decays indicate that there are only three flavors of light neutrino, a fourth family of fermions might have a heavy neutrino, which could be explained, for example, by the seesaw mechanism [20].

A higher cross section than predicted in the s-channel mode of production could indicate the presence of other processes that have the same final state. Besides WH Higgs production, there are many non-Standard-Model processes that could cause an enhancement in this channel, such as production of a heavy right-handed W' boson, strong dynamics, production of a charged top pion, or Kaluza-Klein modes of the W boson. Other theories predict enhancement in the t-channel, including anomalous flavor-changing neutral currents. Several other theories have been proposed in the literature [21].

Chapter 2

Experimental apparatus

Because of their remarkable mass, top quarks cannot be produced by most particle accelerators, which lack the energy. Fixed-target experiments are impractical for this purpose; particle colliders, which provide much more center-of-mass energy, are necessary [22]. A circular synchrotron in the most practical type of collider, since it minimizes the amount of tunnel and number of acceleration stations needed to reach high energies.

Besides the large energies required, another factor to consider is the cross-section of top quark production. The large mass of the top quark makes it kinematically more difficult to produce than most particles, making the overall cross section much smaller than that of lighter particles. Thus, a very large number of collisions is required to collect enough data to make measurements of top quarks. This requires not only high energy but also high *luminosity*, which is a measure of the rate of particle collisions. Knowledge of the luminosity allows an estimation of the expected number of events N of a given process with a cross section σ.

$$N_{proc} = \sigma_{proc} \cdot \int L \, dt$$

The luminosity of a collider is given by

$$L = \frac{N_p N_{\bar{p}} B f_0}{4\pi\sigma^2}$$

for B bunches of particles, each having N protons or antiprotons colliding at a frequency f_0 with a beam width of σ [23]. Colliders can improve their luminosity by increasing the number of particles, increasing the collision frequency, or decreasing the width of the beam. Experiments measure the size of their data sets by the time integral of the luminosity, which is called *integrated luminosity* and is given in units of inverse cross section, typically pb^{-1} or fb^{-1}.

This analysis is performed at the Collider Detector at Fermilab (CDF), a multipurpose particle detector at the Tevatron, a proton-antiproton synchrotron with a center-of-mass energy of 1.96 TeV at Fermilab in Batavia, Illinois. The Tevatron is the world's highest-energy operating particle accelerator and the only one capable of directly producing top quarks, although its energy and luminosity will soon be eclipsed by the commissioning of the Large Hadron Collider (LHC) [24] at the European Organization for Nuclear Research (CERN) [25] facility in Geneva, Switzerland.

2.1 The Fermilab accelerator complex

A collider's design is vastly simplified if it collides particles with their antiparticles because the particles can travel opposite directions in the same beampipe and be bent by the same set of magnets. The Tevatron collides protons and antiprotons. Although electron-positron collisions are easier to analyze (because electrons are single pointlike particles,

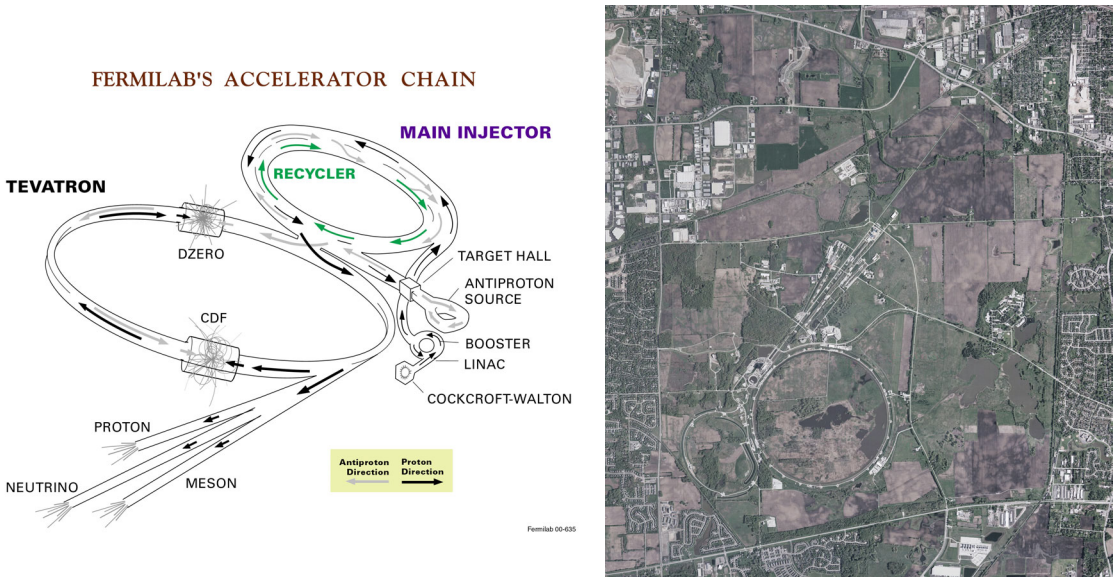

Figure 4. A diagram (left) and aerial photograph (right) of the Fermilab accelerator chain, culminating in the Tevatron, which collides protons and antiprotons at two interaction points, CDF and D0.

unlike the composite protons), the large synchrotron radiation of electrons prohibits their use at high energies in circular accelerators. Because synchrotron radiation increases as the inverse of the fourth power of a particle's mass [26], protons, which have roughly 200 times the mass of electrons, radiate much less. Protons and antiprotons are currently the only viable alternative to electron and positrons.

Protons are, of course, abundant and readily available in nature; antiprotons must be produced and stored. In addition, a single accelerator cannot bring particles from rest to very high energies because no magnets have the dynamic range necessary. Consideration of these requirements led to the design of a chain of accelerators at Fermilab.

The Tevatron [27] is the last in a chain of accelerators that gradually increase the energy of protons and antiprotons. An overview of the entire complex is shown in Figure 4. The Tevatron's first physics run, referred to as Run I, occurred from 1992–1996. After a series of upgrades, it began running again (Run II) in 2002 and is currently still running.

Proton source

The protons used in the Tevatron are originally extracted from very pure hydrogen gas. For ease of insertion into the Booster, the particles accelerated are actually H^- ions instead of protons. Hydrogen gas is moved between two electrodes and a spark ionizes the hydrogen into electrons and H^+ ions. The positive ions strike a cathode made of cesium, which has a low work function and thus loses electrons easily, and occasionally pick up two electrons and form H^- ions. An electrostatic extractor sends them to the preaccelerator.

Preaccelerator

The preaccelerator is a Cockroft-Walton-style [28] electrostatic accelerator. Ions from the proton source are subjected to a potential of –750 kV, thus producing beams of H^- ions with an energy of 750 keV. The H^- ions are steered and focused by magnets down a transfer line to the Linac.

Linac

The next stage of acceleration is a linear accelerator, abbreviated Linac, built in two sections. The older section consists of five *drift tubes*, modeled after Luis Alvarez' original proton linear accelerator [29], that accelerate the ion beam to 117 MeV. The newer section,

added in 1993, has seven side-coupled cavity modules that accelerate the H^- ions to 400 MeV. Both sections work on the same basic principle of acceleration.

Drift tubes use an alternating-current electric field (usually referred to as *RF*, for "radio frequency") and are designed so that particles are shielded from the field when it points opposite their direction of motion. Thus, the H^- ions only feel a force from the electric field when it will increase their speed. Because the particles are accelerating, the cavities must become gradually longer along the length of the Linac. This RF method of acceleration, used by all accelerators at the Fermilab complex, causes the particles to group together into *bunches*, and from this point on most particle beams are characterized by an RF bunch structure.

The side-coupled modules in the Linac have an accelerating gradient of 7.5 MV/m, three times that of the drift tube modules. Some protons are diverted after the fourth module to hit a beryllium target to generate neutrons for use at the Neutron Therapy Facility [30]. At the far end of the Linac is a "chopper" that electrostatically selects a portion of the Linac beam to be sent along a transfer line to the Booster. The Linac completes fifteen acceleration cycles per second.

Booster

The Booster is a proton synchrotron, approximately 150 meters in diameter, that accelerates protons to 8 GeV. It has the same duty cycle as the Linac, 15 Hz. The acceleration is accomplished by eighteen ferrite-tuned RF cavities located around the ring.

Ninety-six conventional magnets with a maximum field of 0.7 T bend the beam into a circular orbit. The Booster is able to hold multiple batches of particles from the Linac at once to increase beam intensities, often storing eleven or twelve batches in its ring.

A special set of magnets handles the injection of incoming H^- ions from the Linac. Magnets bend the circulating protons and the injected H^- ions into a single beam that points through a sheet of foil. The foil strips the electrons from the H^- ions, leaving behind only protons. A similar set of magnets steers the beam back into the Booster orbit while removing any leftover H^- ions.

A set of fast kicker magnets extracts the proton beam from the Booster. Some of the protons are sent to the Mini Booster Neutrino Experiment (MiniBooNE) [31]; the rest go into a transfer line that leads to the Main Injector.

Main Injector

The Main Injector is a large proton synchrotron with a diameter of about 1 km. It has two main functions involving the Tevatron: accelerating protons and antiprotons to 150 GeV for injection into the Tevatron, and accelerating protons to 120 GeV to be sent to the antiproton source. In addition, it sends protons to the Neutrinos at the Main Injector (NuMI) beamline [32] and various fixed-target experiments.

The Main Injector uses 344 dipole magnets and 208 focusing quadrupole magnets, all conventional water-cooled electromagnets, to steer the proton beam. It can accelerate protons to 150 GeV in two seconds. At Fermilab, the Main Injector typically alternates

between different operational modes, so that one cycle may be sent to NuMI, the next to the antiproton source, and so on.

Antiproton source

The antiproton source produces antiprotons for use in Tevatron collisions. The Main Injector sends 120-GeV protons down a transfer line to a nickel target. (Nickel was chosen because it can absorb more heat without melting than other metals.) Antiprotons are among the products resulting from this collision; they are selected by an electromagnetic selector and focused down a transfer line to the Debuncher. Studies have shown that 120 GeV is the optimal energy for antiproton production; at this energy, approximately one antiproton is collected per 10^5 protons sent to the antiproton source. The resultant antiprotons have an average energy of about 8 GeV.

Debuncher

The Debuncher is not an accelerator but a triangular storage ring. Its main purpose is to "debunch" the particle beam, removing its RF bunch structure. Magnets in the Debuncher decrease the momentum spread of the antiprotons by rotating them in phase space, trading momentum spread for time spread. This results in a beam of particles that have no RF bunch structure but have roughly uniform momentum. Antiprotons remain in the Debuncher until the next batch of protons is sent to the antiproton target, at which point the antiprotons are sent to the Accumulator.

Accumulator

The Accumulator lies in the same tunnel as the Debuncher. It is a long-term antiproton storage ring, designed to store antiprotons with minimal losses for days. Antiprotons from the Debuncher are manipulated by RF systems in the Accumulator to fill a stable region of phase space, known as the *core*. The core is kept as small as possible to minimize the momentum spread of the antiprotons; a smaller beam gives a higher luminosity upon injection into the Tevatron.

While the antiprotons stay in the Accumulator, they are "cooled," reduced in transverse momentum, through a process called *stochastic cooling* [33]. This procedure measures the momentum spread of a group of antiprotons and sends a signal across the ring to corrector magnets, which adjust their fields for each group of particles to reduce the momentum spread of those particles. This results in denser antiproton beams injected into the Tevatron, increasing the resulting luminosity.

Extraction from the Accumulator requires the antiprotons to be collected into bunches again. Adiabatic activation of RF stations causes a portion of the beam to be collected into bunches, which are then transferred back to the Main Injector, decelerated to 8 GeV, and injected into the Recycler.

Recycler

When the Accumulator reaches its maximum optimal capacity, its antiprotons are passed into the Recycler, a ring of permanent magnets in the same tunnel as the Main

Figure 5. Photographs of components of the Fermilab accelerator complex: (top left) the Cockroft-Walton preaccelerator, (top right) the interior of one of the Linac drift tubes, (center), the Debuncher and Accumulator, (bottom left), the Booster, (bottom center), the Main Injector, and (bottom right) the Tevatron.

Injector. This storage ring keeps antiprotons at 8 GeV, collecting them until the Tevatron is ready for injection. In the Recycler, antiprotons are cooled further using a process called *electron cooling* [34], in which a beam of electrons is accelerated to the same energy as the antiprotons and run alongside it. Transverse momentum from the antiproton beam is passed to the much lighter electrons, causing the antiprotons to lose transverse momentum, making the beam smaller. Antiprotons are injected from the Recycler to the Main Injector, which accelerates them to 150 GeV for injection into the Tevatron.

Tevatron

The Tevatron is a large synchrotron, 1 km in radius, that accelerates particles from 150 GeV to 980 GeV. It keeps both protons and antiprotons in the same beampipe, revolving in opposite directions. Electrostatic separators produce a strong electric field that keeps the two beams from touching except at the collision point. The beam is steered by 774 superconducting dipole magnets and 240 quadrupole magnets with a maximum magnetic field of 4.2 T. They are cooled by liquid helium to 4.2 K, at which point the niobium-titanium alloy in the magnets becomes superconducting.

The Tevatron holds 36 bunches each of protons and antiprotons. The process of injecting particles into the machine, accelerating them, and initiating collisions, referred to as a *shot*, starts with injection of protons, one bunch at a time, at 150 GeV from the Main Injector. The antiprotons are injected four bunches at a time from the Recycler through the Main Injector. RF cavities accelerate the beams to 980 GeV, and then some electrostatic

Figure 6. Photographs of the CDF detector during its installation: (left) the insertion of the silicon detector into the center of the tracking chamber, surrounded by the calorimeter, with the plug pulled out, and (right) the detector as seen from the side, with the muon chambers exposed.

separators switch polarity to cause the beams to collide at two points. Each interaction point lies at the heart of a particle detector: one named D0 (for the technical name of its position in the Tevatron ring) and the other named the Collider Detector at Fermilab (CDF).

2.2 The CDF detector

Particle detectors measure the momentum and energy of particles that result from particle interactions. CDF [35] is a multi-purpose detector used for a broad range of physics analyses and thus has many components used for many different purposes. Most physics analyses, however, use the basic functions of tracking and calorimetry.

When dealing with particle detectors it is conventional to give coordinates in z, the longitudinal position from the beam pipe; φ, the polar angle (where $\varphi = 0$ is parallel to the ground and points out of the Tevatron ring); and the *pseudorapidity*:

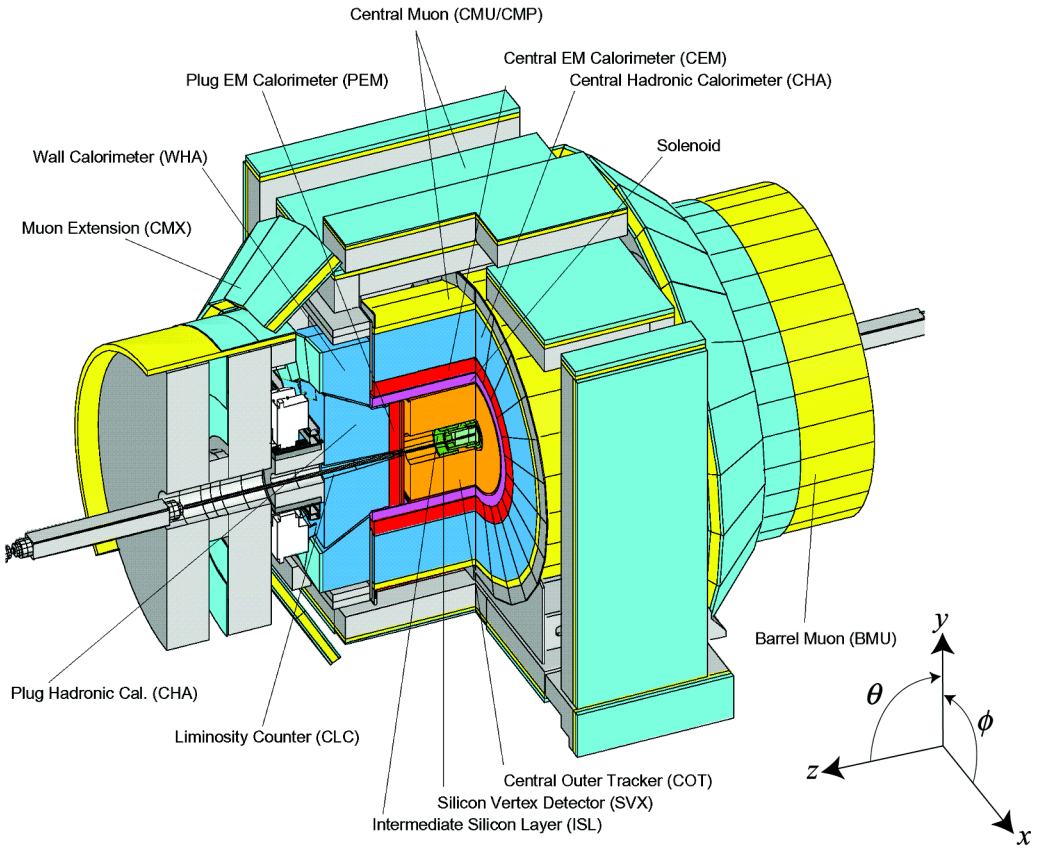

Figure 7. A diagram of CDF with the different subdetectors labeled.

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right)$$

Pseudorapidity is closely related to the *rapidity* of a particle, $y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right)$, which is a useful quantity because differences in rapidity are invariant under Lorentz transformations. Pseudorapidity is a very good approximation for rapidity when the energy of a particle is much larger than its mass (as is nearly always the case at the Tevatron), and it is useful

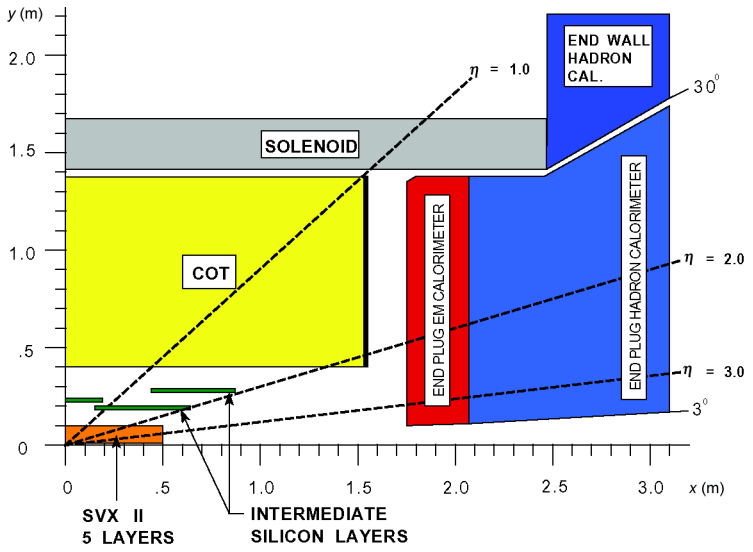

Figure 8. A cut-away view of the tracking volume in the *x*–*y* plane. Lines of constant pseudorapidity (η) are labeled on the picture.

because it is a Lorentz-invariant quantity that describes the azimuthal position of a particle independent of reference frame.

Tracking system

The tracking system is used to reconstruct the trajectories of charged particles. These detectors are placed inside a magnetic field to allow a measurement of the momentum from the curvature of the track (and, from the direction of the curvature, the charge of the particle). Because resolution is very important in distinguishing the tracks of many particles, the tracking detectors are placed close to the interaction point. Good resolution is also important for detecting displaced secondary vertices, a common feature of *B* hadron decays.

At CDF, a silicon detector lies nearest the interaction point, surrounded by a proportional wire chamber.

Silicon detectors

The best tracking resolution comes from the silicon microstrip detectors, placed closest to the interaction point. These are constructed of wafers of p-type silicon with thin (100-μm) strips doped with n-type silicon. A reverse bias voltage extends the depletion region to include the entire strip of silicon. When a charged particle strikes a strip, it ionizes the silicon, creating electron-hole pairs. The voltage draws electrons to one end of the strip, the holes to the other end, and the resulting charge is collected by readout chips mounted at the end of the sensors [36]. Silicon detectors have the advantage of excellent resolution; however, they are expensive and sensitive to radiation damage.

Most of the silicon at CDF is double-sided, with the strips on the two sides perpendicular to each other, allowing for a simultaneous two-dimensional measurement of a particle's position. Multiple layers of this silicon makes three-dimensional reconstruction of particle tracks possible. In total, the silicon detectors form a cylinder 1.9 m long with a radius of 28 cm. The silicon system is divided into three subdetectors [37].

Layer 00: The innermost silicon subdetector is a layer of single-sided silicon attached directly to the beampipe called Layer 00 (so named because it was added after the innermost layer of the Silicon Vertex Detector, which was called Layer 0). The

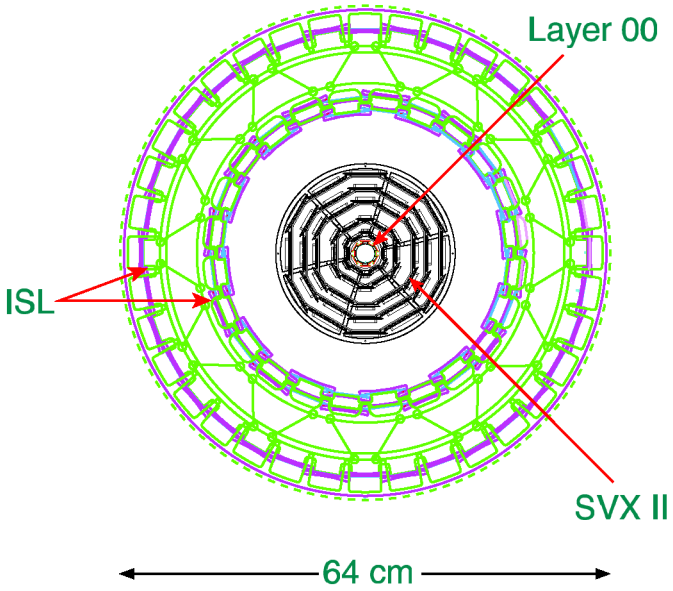

Figure 9. A frontal view of the silicon subdetectors.

strips of silicon lie only 1.6 cm from the beamline and improve the resolution of the position of secondary vertices.

Silicon Vertex Detector (SVX II): The main part of the silicon detector is a set of five concentric layers of double-sided silicon that are distributed around the beampipe at radii from 2.1 cm to 17.3 cm. This detector provides high-resolution tracking information and is particularly useful for resolving displaced secondary vertices.

Intermediate Silicon Layers (ISL): The ISL serves two purposes: to help follow tracks between the Central Outer Tracker and the SVX, and to extend coverage of tracks to a pseudorapidity of $|\eta| < 2.0$. This improves the performance of silicon-only forward tracking and improves track resolution in the central region.

Central Outer Tracker (COT)

The COT [38] is a cylindrical drift chamber that is the workhorse of the tracking system. Eight radial sections, or *superlayers*, of wires are strung through a volume containing a mixture of argon and ethane gasses. When voltage is applied to the wires, a charged particle flying through the chamber ionizes the gas, and the differently charged ions are attracted to the wires at different electric potentials. The charge collected by the wires is collected at readout boards at the end of the chamber to determine which wires were near a charged particle, and the time difference between arriving pulses gives information on the longitudinal position of the particle [39]. This style of drift chamber is relatively inexpensive and has relatively fast readout, allowing the trigger system to reconstruct tracks very quickly. The eight superlayers of the COT are arranged in alternating axial and stereo sections: axial superlayers lie parallel to the beamline, while stereo superlayers have a 2° offset from the parallel. These provide information on the ϕ and z position, respectively, of a track.

Time-of-Flight system (TOF)

Timing information can be important for identifying particles. Measuring the time it takes for a collision product to reach a given part of the detector can be used to discriminate between particles of different masses, such as pions and kaons. This analysis, though it does not use this discrimination directly, uses timing information to recognize and remove events coming from cosmic rays. The TOF system is useful for this purpose.

The TOF system [40] is a ring of scintillators that surrounds the COT. Charged particles passing through them create bursts of light which are collected by *photomultiplier tubes* (PMTs). PMTs take advantage of the photoelectric effect by amplifying the electrons released when photons strike a metal plate, converting light into a proportional electrical signal which is sent to readout electronics. Because PMTs perform less well in a strong magnetic field, an additional preamplifier is needed to boost the signal. The TOF system measures the difference in time between the time of a particle interaction and the arrival of a charged particle. It has a resolution of about 100 ps.

Solenoid

The tracking chambers are contained within a large superconducting solenoidal magnet [41] that produces a magnetic field of 1.4 T parallel to the beamline. It is made of aluminum-stabilized niobium-titanium and runs at a current of about 4650 A. Charged particles bend in this field and the curvature of the reconstructed tracks allows a measurement of their momentum.

Calorimeters

In addition to tracking information, which determines the trajectory and momentum of a particle, detectors also must measure a particle's energy. This is done in the calorimeter system, a series of layered detectors that measures the energy deposited by particles passing through it. A particle passing through a material will lose energy through ionization and absorption (and *Bremsstrahlung*, for electrons). Collecting and measuring the radiated

energy allows an estimation of the particle's original energy. The calorimeters at CDF use plastic scintillators to collect energy from a charged particle. Energetic particles passing through a scintillator excite the atoms in it, causing light to be emitted. This light is collected and guided by wavelength-shifting fibers to PMTs.

As a charged particle passes through a material, it scatters off the atoms in it, creating *showers* of less energetic particles. These secondary particles scatter to create more particles, so the number of particles in the shower increases. At the same time, the particles lose energy and eventually are lost by ionization, so the number of particles in the shower eventually decreases. The measurement of the full shower profile allows a more precise determination of a particle's energy; thus, most calorimeters have multiple layers.

Because the radiation length of scintillator is large, a large amount of it would be needed to collect all the energy of a particle. CDF reduces the amount needed by alternating layers of scintillator with layers of a dense absorber. This creates a *sampling calorimeter*, which measures a fraction of the total energy of a particle and estimates the true energy based on calibrations with test beams. While this reduces the precision of the energy measurement, it keeps the calorimeter from becoming prohibitively large.

Electrons and photons, having very little mass, radiate their energy much faster than heavier particles. Mesons and hadrons (predominantly pions, kaons, protons, and neutrons) lose far less energy to ionization, but they also lose energy from inelastic nuclear scattering mediated by the strong force. This motivates the construction of two distinct calorimeters:

an *electromagnetic calorimeter* and a *hadronic calorimeter*. Because of the slower rate of radiation loss of hadrons, the hadronic calorimeter must be much larger than the electromagnetic. However, roughly 30% of hadronic energy in hadronic showers is unmeasurable: it is lost in the breakup of nuclei, nuclear excitation, and similar effects [42]. Thus, the energy resolution of the electromagnetic calorimeter is much better than that of the hadronic.

Muons, which are massive but do not interact by the strong force, leave only minimum ionizing energy in the calorimeters. They must be measured separately by muon chambers.

Central calorimeters

The central section of the calorimeter directly surrounds the solenoid and covers a pseudorapidity of $|\eta| < 1.1$. It is arranged in concentric layers that lie parallel to the beam line. The electromagnetic calorimeter is closer to the beam, immediately outside the solenoid; it is surrounded by the hadronic calorimeter [43].

Central Electromagnetic Calorimeter (CEM): The electromagnetic calorimeter is composed of 31 alternating layers of lead and scintillator. It is segmented into a series of "towers"—24 towers in ϕ and 10 towers in η—which determines the granularity of the energy measurement. The granularity is important because clusters of energy measured in the calorimeter often must be matched to the track of a

particle in the tracking volume, so finer granularity improves the resolution of particle identification.

Central Electromagnetic Shower Maximum Detector (CES): The point at which the shower has the greatest number of particles is called the *shower maximum*, and its average position is six radiation lengths into the electromagnetic calorimeter. A special layer of the calorimeter is inserted at this point, more finely sectioned than the surrounding layers, to allow for better position resolution of the calorimeter clusters. The CES is a series of strip and wire chambers that measures the position of the charged particle to much better precision than the coarsely segmented calorimeter, thus improving cluster position resolution and making it easier to match clusters with tracks from the tracking detectors.

Central Hadronic Calorimeter (CHA): The hadronic calorimeter is composed of 32 alternating layers of iron and scintillator. It has the same polar segmentation as the CEM but is only eight towers long in pseudorapidity, since due to its geometry it only covers up to a pseudorapidity of $|\eta| < 0.9$.

Wall Hadronic Calorimeter (WHA): The central section of CDF is cylindrical. This means that the hadronic calorimeter cannot extend as far forward as the electromagnetic calorimeter, which leaves a gap in the calorimeter. This gap is filled by the WHA, which is a section of 15 layers of alternating iron and scintillator attached to the longitudinal face of the detector. It has similar construction to the

CHA but its layers are oriented perpendicular to the beam. It has full coverage in ϕ and six towers in η to cover a pseudorapidity of $0.8 < |\eta| < 1.2$.

The hadronic calorimeters have time-to-digital converter (TDC) cards in their readout electronics in order to measure the time of the arrival of particles to the detector. This allows rejection of out-of-time particles resulting from stray beam particles or cosmic rays.

Plug calorimeters

The plug calorimeter [44], so named because it fits into the end of the central CDF detector like a giant plug, is a series of calorimeter layers that are perpendicular to the beam line and allow measurement of energetic particles with a pseudorapidity of $1.1 < |\eta| < 3.6$. The layers of the plug calorimeter are oriented perpendicular to the beam, with the electromagnetic calorimeter closer to the interaction point.

Plug Electromagnetic Calorimeter (PEM): The electromagnetic calorimeter is a series of 23 layers of alternating lead and scintillator. It is segmented into 12 tower groups in η. The inner four groups have 24 towers each in ϕ; the outer groups have 48.

Plug Electromagnetic Shower Maximum Detector (PES): The PES detector is based on the same principle as the CES detector: it is a calorimeter with finer position resolution to improve the precision of the position of a calorimeter cluster and it is located six radiation lengths into the PEM. The PES detector is composed of a series

of scintillator strips. Each octant of the calorimeter has two layers of two hundred strips each, oriented at a relative 45° angle.

Plug Hadronic Calorimeter (PHA): The hadronic calorimeter is composed of 23 layers of alternating iron and scintillator. Because of the geometry of the plug, it only covers a pseudorapidity of $1.2 < |\eta| < 3.6$ (the remaining region is covered by the WHA). Its segmentation is the same as the PEM, except the PHA has one fewer tower group in η.

Muon chambers

Muons require special treatment in a particle detector. Being relatively long-lived, they are one of the few unstable particles to make it through the detector without decaying; being large in mass, they deposit only minimum ionizing energy in the electromagnetic calorimeter; and being leptons, they leave very little energy in the hadronic calorimeter. This makes them very difficult to identify from tracking and calorimetry information alone.

Muon chambers are mounted outside the calorimeters to identify muons. They rely on the observation that few other particles make it through the calorimeter without losing most of their energy. Some muon chambers are mounted behind thick pieces of steel, decreasing the rate of non-muonic particles. The muon chambers at CDF [45] are a series of single-wire drift chambers that detect the charged muons passing through them and register the result as a muon *stub*. A stub requires a hit in three of the four layers of drift

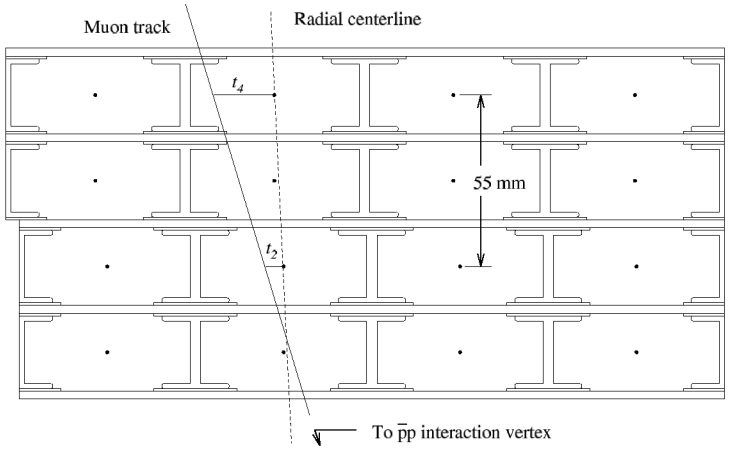

Figure 10. A cross-sectional view of a muon chamber, showing the path of a charged particle passing through the detector.

chambers. When a muon stub is matched to a track in the tracking subdetectors, its energy can be computed from the known mass of the muon and the measured momentum.

Because muon stubs can also be generated by cosmic ray muons, timing information helps identify muons that come from collisions. Scintillators paired with the drift chambers provide timing information that is used to reduce the cosmic ray background. Because the muon chambers are the outermost components of the detector, they are most constrained by the size and geometry of the collision hall, sometimes resulting in unusual configurations.

Central Muon Detector (CMU): The CMU detector is built directly outside the CHA calorimeter. It covers the central region up to a pseudorapidity of $|\eta| < 0.6$. The CMU contains four radial layers of drift chambers to reconstruct the track of a muon as it passes through them.

Central Muon Upgrade (CMP): One downside of the CMU detector is its lack of shielding, which makes it easier for other particles to fake a muon signal. The CMP addresses this problem by using large pieces of steel (the return yoke for the solenoid magnet) to absorb other particles, improving the overall trigger efficiency when combined with the CMU. Because it uses the return yoke, the CMP is constructed as a box rather than a cylinder.

Central Muon Extension (CMX) and scintillators (CSX): The CMX subdetector extends the coverage of the muon chambers to a pseudorapidity of $0.6 < |\eta| < 1.0$. Because the collision hall was not designed with this subdetector in mind, the CMX needs several pieces to provide full angular coverage: the arches, which fit in on the sides; the keystone, which sits on top of the detector; and the miniskirt, which goes beneath the detector. These three pieces provide full angular coverage for the detector. The CSX scintillators are used for timing information to reject events that are not consistent with a muon arriving from the interaction point.

Barrel Muon Detector (BMU) and scintillators (BSU and TSU): The BMU subdetector extends the pseudorapidity coverage to $1.0 < |\eta| < 1.5$. These drift chambers and scintillators are attached to the large cylindrical pieces of steel that are attached to the ends of the detector to shield the electronics in the collision hall from beam radiation. The drift chambers form a cylinder (or barrel; hence the name) around

these pieces of steel. Scintillators mounted in gaps in the toroidal pieces of steel make up the TSU and provide additional information for rejecting cosmic ray events.

Čerenkov Luminosity Counters (CLC)

Measurement of acquired luminosity is an essential part of determining the sensitivity of a measurement and estimating the number of background events. CDF has luminosity counters mounted very close to the beamline in the plug, at a pseudorapidity of $3.7 < |\eta| < 4.7$. The CLC [46] is filled with isobutane and designed to detect the burst of Čerenkov radiation that results from a charged particle flying through it. The light is collected and sent to a PMT to be read out. There are three layers in the CLC, with 16 counters in each layer on each side.

Data acquisition system

Every time bunches of protons and antiprotons collide at CDF, the data acquisition system sends a trigger to the subdetectors in the system, telling them to take a measurement. However, with collisions occurring every 396 ns, it is impossible for every event to be recorded to disk. It is unnecessary as well; of all the events that occur at CDF, very few are of interest to physicists. Selecting only those events which are of interest to physicists is the job of a fast *trigger* system [47]. This allows CDF to keep the events it needs and still operate under practically deadtime-free conditions.

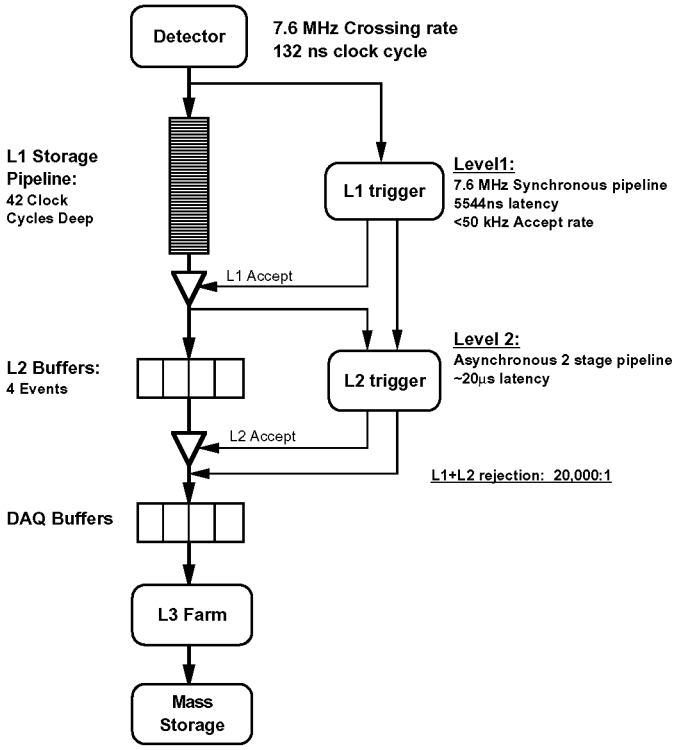

Figure 11. A functional block diagram of the CDF data acquisition system.

Level 1

The lowest level of trigger at CDF must make a decision within 5 μs of each collision. Collisions occur at a rate of 1.7 MHz; the Level 1 trigger system reduces the acceptance rate to about 40 kHz. Because of the stringent time requirement, it is implemented in hardware. The Level 1 trigger can make its decision based on clusters of energy in the calorimeters (from jets or photons), electrons and muons (tracks matched to muon stubs or calorimeter clusters), undetected energy inferred from conservation laws, or the sum of calorimeter energy. A system called the Extremely Fast Tracker (XFT) [48]

reconstructs tracks in the COT. A dedicated processor makes rapid decisions based on these quantities.

Level 2

The next level of trigger looks at the events that passed Level 1 and refines the selection, making a decision within a comparatively long 30 μs. The extra time allows it to perform more careful clustering in the calorimeter, including information from the CES; it also uses silicon information. Because of the large number of readout channels, the silicon readout chips hold each event on an onboard circular buffer. They do not send the information to the data acquisition system until an event passes the Level 1 trigger.

Silicon information is processed by the Silicon Vertex Tracker (SVT) [49], a system that rapidly analyzes the silicon data to look for a displaced vertex. The calculations of the different systems (calorimeter, COT, and silicon) are sent to a single computer that analyzes the data and decides whether to accept the event. The Level 2 system is asynchronous—it does not require an event to be finished by a fixed time after the collision—and it reduces the rate of accepted events to about 400 Hz.

Level 3

The highest level of trigger at CDF is implemented in software on a farm of several hundred computers. Each event that passes Level 2 is sent to the event builder, which assembles the disparate information into a data format readable by Level 3. Level 3 fully reconstructs the event and analyzes high-level quantities to make the final acceptance

decision. Events that pass Level 3 are ready to be written to disk. Level 3 reduces the event rate to about 100 Hz.

Prescales

The trigger system includes many different triggers which are used for different analyses and calibration studies. Because some triggers fire at a very high rate, it is necessary to attach a *prescale* to them to keep the overall rate low. This involves rejecting a fixed fraction of the events; for example, a trigger with a prescale of 10 only keeps every tenth event which would otherwise pass.

The rate of a trigger increases with the instantaneous luminosity of the collisions. Thus, as the luminosity declines, more bandwidth is available in the trigger system, so it becomes useful to use *dynamic prescaling*, which changes the prescale on the triggers as the instantaneous luminosity changes. Triggers with dynamic prescales require careful bookkeeping to properly account for the luminosity they record.

Consumer Server/Logger (CSL)

Once an event is accepted by the trigger, is needs to be recorded. This is managed by the CSL, which categorizes events by the triggers they fired and writes them to hard disk, reserving a fraction to be used for online monitoring. The data on these disks are then copied to tape storage, ready to be processed with offline reconstruction algorithms for use by physicists.

Chapter 3

Monte Carlo simulation

A complicated analysis with a particle detector requires a good simulation. It is essential to be able to simulate the interaction of particles and the detector in order to understand the behavior of different physics processes. This requires the generation of *Monte Carlo* events, randomly generated events that simulate different physics processes. (The name comes from a city famous for its gambling [50].) There are several steps in creating Monte Carlo events: generation of the initial-state partons, generation of the parton-level final state, showering and hadronization of the products, and simulation of the behavior of these particles in the detector.

3.1 Parton distribution functions

A given physics process at a hadron collider begins with two quarks or gluons in the initial state. These quarks and gluons come from protons and antiprotons, either as valence quarks or extracted from the sea of virtual particles. This means that the initial particles in a collision may be a quark and an antiquark, a quark and a gluon, or two gluons; they may be up or down, charm or strange, or bottom quarks. (Theoretically, there could be top quarks as well, but their mass is so large that the chance of producing them from the quark sea is essentially zero.) The momentum distribution of the constituent *partons* (a general term

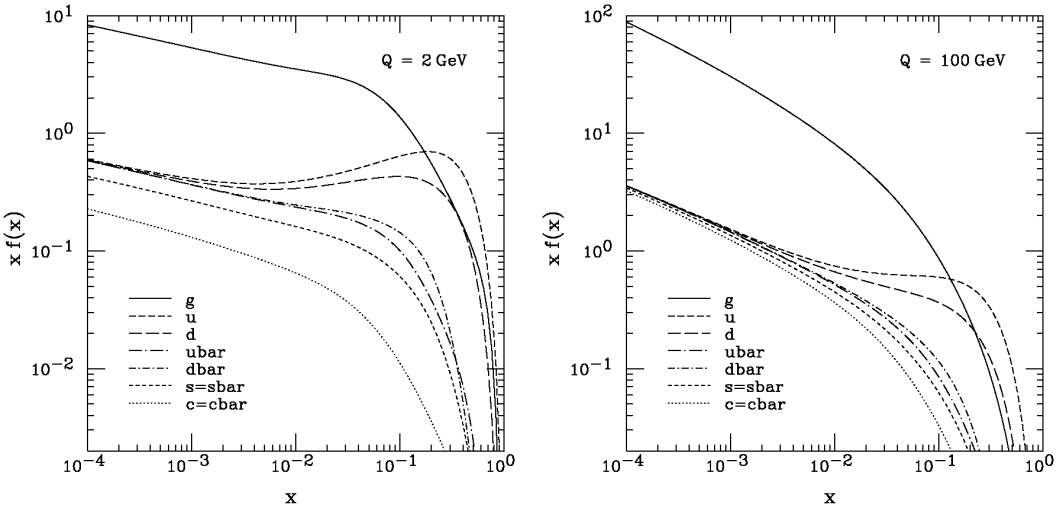

Figure 12. Next-to-leading-order PDFs for protons as a function of momentum fraction (x) at a momentum transfer Q of 2 GeV (left) and 100 GeV (right), from the CTEQ6M parameterization.

which includes quarks and gluons) in a proton is given by *parton distribution functions* (PDFs).

PDFs are determined for each flavor of quark and antiquark, as well as for gluons, in a proton. They give the probability density for finding a parton with a given fraction of longitudinal momentum in an interaction with a given momentum transfer Q^2. An event generator starts with a pair of partons with a certain momentum, and assigns weights to each event based on the PDFs. Because they rely on non-perturbative QCD effects, PDFs require input from experimental data. This analysis uses leading-order PDFs calculated by the Coordinated Theoretical-Experimental Project on QCD (CTEQ), updated in 2002 to include results from QCD studies at HERA and the Tevatron [51] (Figure 12).

3.2 Event generation

The calculation of a hard scattering process is based on tree-level Feynman diagrams and is, in many cases, straightforward. In most cases, the simplest leading-order diagram is calculated as a hard scattering process and radiated leptons, photons, or gluons are treated by a showering algorithm. Next-to-leading-order Monte Carlo generators are starting to become available [52], but have not yet been sufficiently validated on CDF data to use in this analysis.

A Monte Carlo generator uses an "unweighting" method to simulate the relative rate of different event kinematics [53]. First, it creates a large number of events with randomly assigned kinematic properties. It calculates a weight for each event based on the differential cross section for the event's kinematic properties. Then it converts each weight to a probability, taking the highest weight to be unity. This gives the relative contribution of each region of phase space. The generator then examines each event again, choosing a random number between 0 and 1 for each event and keeping only events for which the random number is less than the probability for that event. This results in a set of discrete, unit-weight events whose kinematics, for a large number of events, properly reflect the differential cross section of the process.

PYTHIA

The most convenient event generator to use is called PYTHIA [54], in honor of the oracle of Apollo at Delphi, who would declare the will of the gods—but whose declarations were difficult to interpret. This program contains showering routines as well as an event generator and it takes little effort to pass events between the two. The event generator in PYTHIA can handle simple Feynman diagrams; however, it does not include spin correlations of polarized top quarks, and it uses a parton shower approximation to account for the effects of initial- and final-state radiation which does not include color information. When color and polarization effect are not significant, however, it performs very well.

MadEvent

MadEvent [55] is a Monte Carlo generator that can calculate arbitrary tree-level diagrams with full color and spin polarization information included. It is used for diagrams in which the polarization of the top quark is an important part of the event kinematics.

ALPGEN

Processes with an electroweak boson and radiated gluons are difficult to deal with because of the large amount of radiation they produce; the showering approximation used by PYTHIA, being based only on the tree-level diagram, does not include effects of color flow. However, a full calculation of the matrix elements involved is difficult because the number of distinct diagrams grows as the factorial of the number of jets.

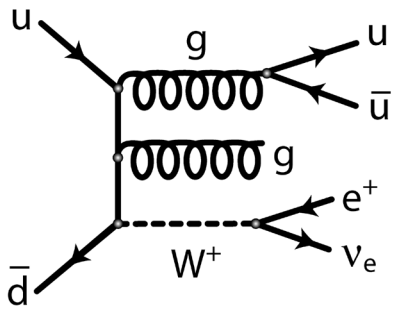

Figure 13. An example of a diagram calculated by ALPGEN. The radiated gluons and quarks are calculated directly from the matrix element by ALPGEN, whereas PYTHIA uses a parton-shower approximation.

ALPGEN [56] is a Monte Carlo generator designed specifically for processes whose final state contains an electroweak boson and several radiated quarks and gluons, a major background this analysis. ALPGEN calculates the matrix elements for processes with gluon radiation and passes the color information to the showering algorithm. This should give a more accurate modeling of the kinematics of the process than PYTHIA's showering approximation, since it includes proper matrix element calculations of the event. ALPGEN also calculates the leading-order cross section of each interaction it generates, which is useful for combining different processes.

3.3 Parton showering

All events, regardless of how they were generated, are passed to PYTHIA for parton showering [57]. This procedure generates initial- and final-state gluon radiation for each

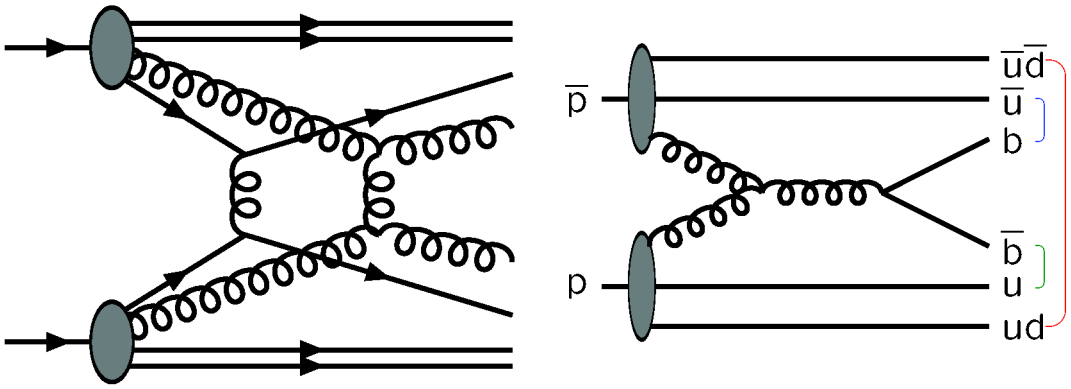

Figure 14. Illustration of (left) multiple parton interactions and (right) beam remnants.

event and allows them to decay to quark pairs, increasing the number of particles in the final state of the event. More particles may be added from effects of beam remnants or multiple interactions (Figure 14). This gives the final set of particles that are passed to the hadronization routine.

PYTHIA generates showers based on probabilities of a particle splitting into two others. It orders the particles by mass and then randomly decides which ones will split, based on the probability of the occurrence. It continues splitting the products until they fall below a pre-set energy threshold. The advantage of this method is its speed and its ability to generate an arbitrarily large number of jets without needing to recalculate matrix elements. On the other hand, it works only as an approximation and lacks color flow information from the matrix elements. However, color flow is not a major concern for electroweak processes because they have only one gluon in next-to-leading-order diagrams;

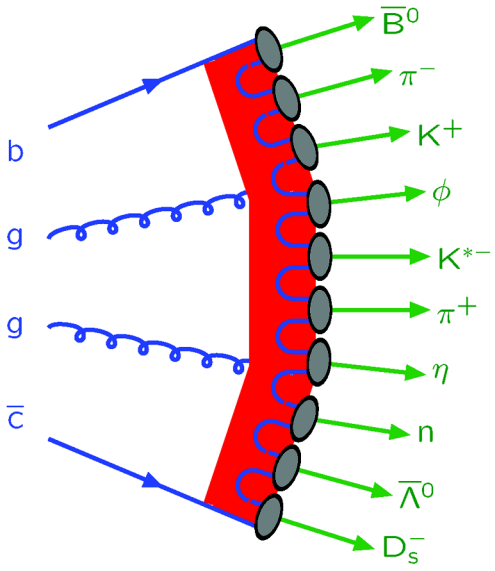

Figure 15. An illustration of the hadronization process preformed by PYTHIA.

only for diagrams with many gluons is it a major concern, requiring ALPGEN for a proper calculation.

3.4 Hadronization

Once the final-state particles have been generated, it is necessary to hadronize the quarks and gluons—to create quark-antiquark pairs that form baryons and mesons, since bare quarks and gluons are not observable. This step causes quarks and gluons to form into jets of hadrons which are measured in the detector.

PYTHIA performs its hadronization using a color string model. Each pair of quarks is modeled as though connected by a relativistic string which increases linearly in energy as

separation increases. As the energy in the string increases, it becomes increasingly more probable that it will form a new quark-antiquark pair. These new particles can then be grouped with the original ones to form mesons and baryons. The new quarks are produced with a flavor ratio, based on experimental data, of $u : d : s = 1 : 1 : 0.3$. Heavier quarks are assumed not to be created in the hadronization process.

Most of the particles resulting from hadronization are unstable, so PYTHIA causes them to decay into relatively stable particles (electrons and muons, protons and neutrons, pions and kaons) that can actually be detected. This step uses branching ratios and lifetimes measured in various experiments to calculate the final decay products. In this procedure, PYTHIA ignores spin information and uses a simplified algorithm for B mesons and tau leptons, so it must be supplemented with separate algorithms for these cases.

For the decay of hadrons that include bottom quarks, the program QQ [58] is used. This is a Monte Carlo generator written at the CLEO [59] experiment and designed to deal properly with the decay of hadrons with bottom quarks. The decay of tau leptons is performed by the TAUOLA [60] package, which simulates tau lepton decays with full spin correlations, final-state neutrinos, and resonant distributions of intermediate particles.

3.5 Detector simulation

Once the final long-lived particles have been generated, it is important to determine how the detector will respond to them. This requires a full detector simulation which

simulates the response of the different subcomponents of the detector, including resolution effects, inherent inefficiencies in the detector, and the behavior of the particles as they pass through passive material (such as cables or support structures) in the detector. When this is done, the Monte Carlo events can be put into a data structure identical to that obtained from collision data, thus allowing reconstruction algorithms to work exactly the same way on data and Monte Carlo events.

CDF uses a program called GEANT [61] to model the tracking volume of the detector. GEANT allows the construction of a mathematical model of the detector which can simulate the passage of charged particles through it, including showering to secondary or tertiary particles. This is used along with charge deposition models to simulate the response of the tracking detectors (silicon and COT).

By the time a charged particle reaches the calorimeter, it has showered into many secondary particles. Because modeling the interactions of each particle and all its secondary particles is computationally intensive, CDF stops using GEANT after the first inelastic collision occurs in the calorimeter. Instead it switches to a parameterized calorimeter response, tuned to test beam data, which employs a program called GFLASH [62]. This rapidly and accurately simulates the response of the calorimeter towers to the energy deposited by the incoming charged particles, completing the detector simulation.

Chapter 4

Event reconstruction

Once detector data is obtained, either from real particle collisions or Monte Carlo events, it needs to be converted from the raw data in the detector to reconstructed physics quantities. This happens in two general steps: first, information from subdetectors is combined to form high-level detector objects: tracks in the tracking detectors, clusters in the calorimeters. Then these objects are analyzed to associate them with physical objects: electrons, muons, jets, or neutrinos. These can then be used in a physics analysis.

4.1 High-level detector objects

This analysis is primarily concerned with *tracks*, sequences of hits left by charged particles as they pass through the tracking detectors, and calorimeter *clusters*, collections of towers in which energy from particles has been deposited. Associating these together, with quality cuts, allows reconstruction of electrons, muons, and jets and a calculation of their energy and momentum.

Tracking

Most tracking at CDF begins with the COT tracking chamber. First, hits in the COT are identified. These are associated together within each superlayer to form short track

segments in each superlayer, which can then be combined across superlayers to form a track. Because the superlayers of the COT alternate in orientation between axial and stereo superlayers, the tracking algorithm compares segments in all axial superlayers first, starting with tracks in the outermost superlayer and finding the segment that gives the best fit. It then adds the stereo superlayers and performs the fit again to create a final track.

A COT track can be improved by attaching high-resolution tracking information from the silicon detector. The primary algorithm to do this, called Outside-In (OI) tracking, starts with COT tracks and extends them by adding hits in the silicon detector.

Normal COT tracks are required to have hits in at least four of the eight superlayers, so tracking in the forward region requires special treatment. One algorithm, called Inside-Out (IO) first finds tracks in the silicon detector, requiring at hits in at least three layers, and extends them by adding hits in the COT that are not already associated with another track. Another algorithm, called Phoenix, uses a cluster in the plug calorimeter and the primary vertex as two points of the track and looks for hits in between that would complete the track. It uses the energy of the cluster to estimate the momentum of the particle, giving a curvature estimate which it uses to search for hits in the tracking region.

Most tracks in this analysis are from a collection which includes all COT tracks with no silicon, OI tracks, and IO tracks. "Track" in this paper refers to any of these tracks. Phoenix tracks are also used, but only for identifying plug electrons.

Vertexing

The point of the hard scattering interaction between the partons in a proton and antiproton is called the *primary vertex*. The location of the vertex is important for understanding the kinematics of the particles that result from it. Because the interaction region of the particle beams has a substantial volume, the knowledge of the position of a specific interaction affects the measured kinematic properties of the particles that result from a collision.

The vertexing algorithm takes a set of high quality tracks that are energetic and have at least three silicon hits. It takes the thirty most energetic tracks and performs a fit to a primary vertex. It then prunes the collection of tracks with a χ^2 cut and repeats the fit with the remaining tracks, repeating this loop until all tracks pass the cut. This gives the final position of the primary vertex.

Calorimeter clustering

Particles passing through the calorimeter leave energy behind in *clusters* of towers. Collecting towers together to form clusters is important to properly identifying particles in the calorimeter.

The basic calorimeter clustering algorithm starts with a *seed* cluster that has an energy larger than a certain threshold. Adjacent towers with energy above a lower threshold are added to complete the cluster. The position of the cluster is defined by the energy-weighted mean of the towers in the cluster, and the total energy is estimated by the sum of

the energies of the towers in the cluster. Once a cluster is defined, the precision of its position can usually be improved by matching it with a cluster in the shower maximum detector, which is constructed with a similar algorithm but which has higher position resolution. This gives the final cluster position.

4.2 Particle identification

Having reconstructed tracks and clusters, it is possible now to identify physical objects. These objects serve as the starting point of physics analyses, although additional selection cuts are usually required to remove incorrect reconstructions of leptons that contribute to the sample. This analysis uses leptons, jets, and missing transverse energy.

Leptons

In CDF, an electron is identified as an isolated track matched to a calorimeter cluster. A muon requires an isolated track matched to a stub in a muon detector. Because they are massive and do not interact by the strong force, muons leave only minimum ionizing energy in the calorimeter as they pass through it; thus, another requirement for a muon is that it leave minimum ionizing energy in the calorimeter. This requirement reduces fake muon signals resulting from energetic particles that make it through the calorimeter.

Jets

A jet is a collection of a large number of different particles that all point in the same direction. Because of the wide variety of particles of different momentum that can make up

a jet, the simplest way to define a jet is a cluster in the calorimeter that is not already associated with a lepton. These clusters, along with the tracks pointing to them, are identified as jets.

Jets are reconstructed beginning with *seed towers*, calorimeter towers with at least 1 GeV of deposited energy. The jet is built up by adding more towers in a *cone* with a given radius. This radius defines a circle in the calorimeter in a coordinate system defined by polar angle φ and pseudorapidity η. After the jet is defined, the centroid of the jet is calculated with an energy-weighted mean, and the jet is reclustered with that point as the center of a new jet cone. This procedure is repeated until the centroid position stops changing significantly. Jets with more than 50% overlapping towers are combined into a single jet; otherwise, overlapping towers are assigned to the jet with the closer centroid.

A wider jet cone includes more final-state particles in the jet, which allows a more accurate energy measurement, but makes it harder to distinguish jets that are close together. This analysis uses a cone radius of 0.4, which strikes a balance between the advantages of larger and smaller cones.

Jet energy scale corrections

Of all physics quantities CDF measures, the energy of the jets is one of the most difficult. Jets contain a wide variety of particles, including neutrons, which deposit little energy in the electromagnetic calorimeter, and neutrinos, which deposit no energy at all. In addition, some particles leave energy outside the cone used to identify the jet, and the

response of the detector to these particles varies in different regions due to its construction. To estimate the energy of a parton from the energy of its reconstructed jet, a series of jet energy scale corrections must be applied [63].

Pseudorapidity-dependent correction

First, the non-uniformity of the detector in pseudorapidity is accounted for. This results from the difference in clustering performance between the central and plug calorimeters, and from inefficiencies due to cracks between sections of the calorimeter.

This correction is calibrated with dijet events with a transverse momentum of at least 5 GeV above the trigger thresholds and with angle between them greater than 2.7 radians, to reduce background from radiation. One jet is required to point to the region where the calorimeter performs the best ($0.2 < |\eta| < 0.6$) and is used to trigger the event; the other is adjusted to balance the transverse momentum of the event. The same calculation is performed in Monte Carlo events. Both data and Monte Carlo events are given pseudorapidity-dependent corrections to make the jet energy response uniform across the detector.

The systematic uncertainty on this correction comes from changing the event requirements, which estimates the effect of a kinematic bias introduced by event selection requirements. The maximum transverse momentum allowed for a third jet and the maximum missing transverse energy are varied, and the change in the answer is taken as a systematic

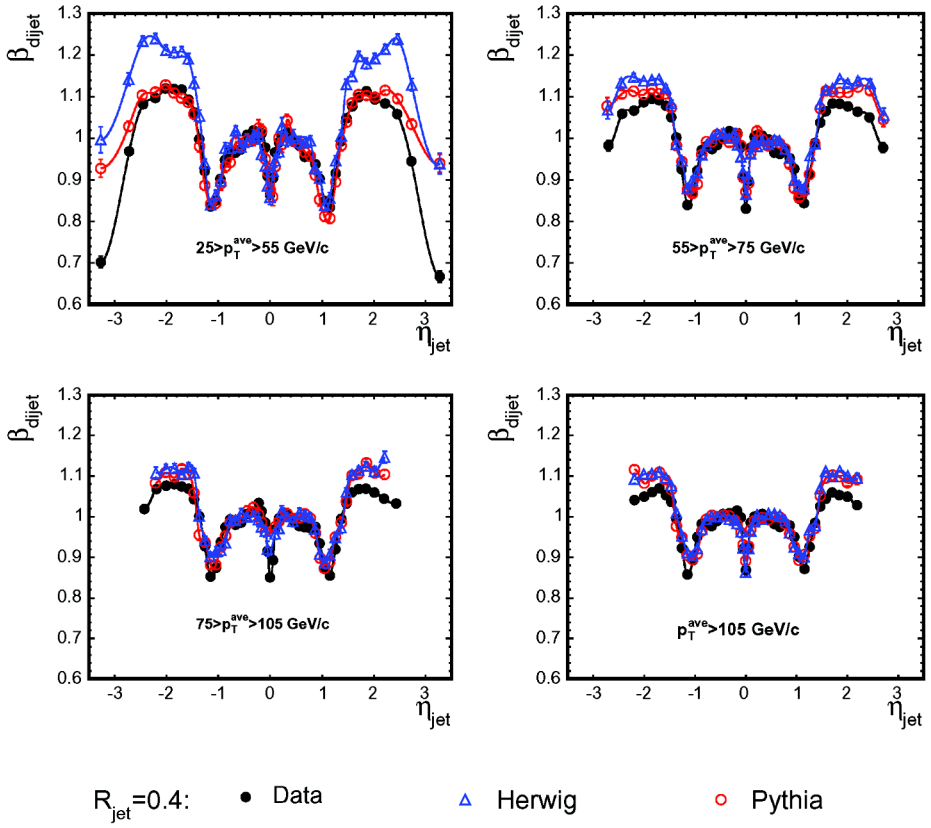

Figure 16. The results of dijet balancing in data and two Monte Carlo generators (HERWIG is not used in this analysis). This shows the variation of the jet energies in pseudorapidity in different regions of transverse momentum. Corrections are applied to remove these variations.

uncertainty. In the central region the uncertainty is about 1%, but at high pseudorapidity and small transverse momentum it can be as large as 7.5%.

Multiple-interaction correction

The next correction deals with the problem of multiple interactions. In general, the number of interactions that occurs when two bunches of protons and antiprotons collide follows a Poisson distribution whose mean increases with instantaneous luminosity. The

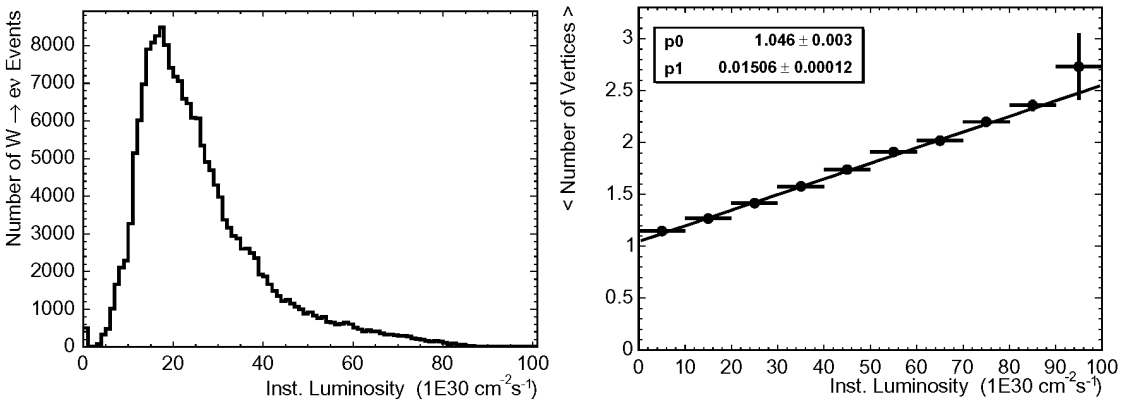

Figure 17. The number of vertices in an event as a function of instantaneous luminosity. These measurements were performed early in the Tevatron's run; in modern running conditions, instantaneous luminosities of up to 300 $cm^{-2}s^{-1}$ are not uncommon.

average number of interactions per bunch crossing varies from one to eight for instantaneous luminosities of 40 to 300 $\mu b^{-1}/s$, which is the approximate range of instantaneous luminosities generated at the Tevatron. These additional interactions cause extra energy to be deposited in the calorimeter, skewing the measured energy.

The number of reconstructed vertices is used to parameterize the number of interactions in an event, since each interaction should have a unique vertex. Minimum bias events, which are triggered by events in the CLC and usually have minimal hard scattering, are used to estimate the effect. A random tower in the best-performing region ($0.2 < |\eta| < 0.6$, which is far from cracks in the calorimeter) of the calorimeter is chosen as a seed tower, a cluster is formed, and its energy is measured. The average energy is calculated for a given number of reconstructed vertices, and the resulting plot is fit to a straight line. The line

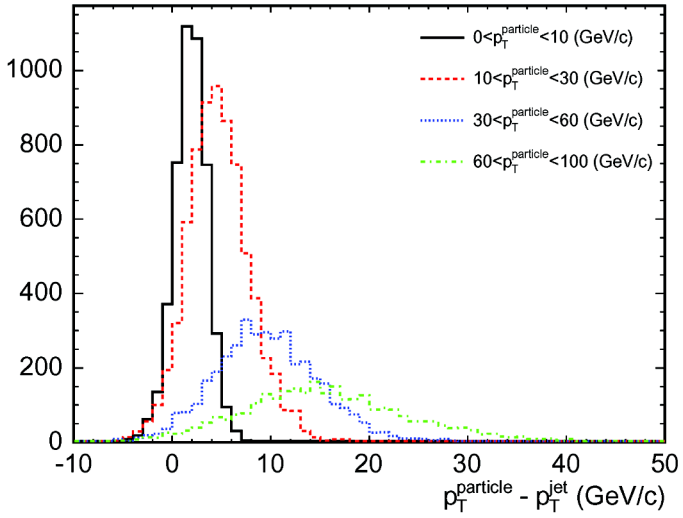

Figure 18. The difference between particle and jet transverse momentum for different particle momenta. This difference is parameterized with a double Gaussian distribution.

derived from this study is used to correct the energy of the jets. This linear approximation works well for the most part but starts to fail with more than seven vertices because of the finite reconstruction efficiency of the vertices.

The uncertainty on this correction comes from performing the same measurement in different samples, including samples with *W* bosons and inclusive jet samples. These account for possible differences in vertex reconstruction efficiency and the rate of improperly constructed vertices. The uncertainty from the difference between these samples is around 15%.

Calorimeter response correction

Another correction adjusts for calorimeter response, allowing a measure of absolute jet energy that can be compared with other experiments. This is derived by creating a Monte Carlo sample of inclusive dijet events in PYTHIA and correcting the jets as already described. The correction is derived from comparing *particle jets,* which are jets at the Monte Carlo level before they are passed through the detector simulation, with *calorimeter jets*, which are reconstructed from the detector simulation. These are required to be within 0.1 of each other in the η–ϕ plane to ensure that they are the same object.

The conversion from particle jet energies to calorimeter jet energies is parameterized by a double Gaussian function in the difference of jet energies $\Delta p_T = p_T^{particle} - p_T^{jet}$:

$$f\left(\Delta p_T\right) = \frac{1}{\sqrt{2\pi}\left(\sigma_1 + N_2\sigma_2\right)}\left(e^{-\frac{1}{2}\left(\frac{\Delta p_T - \mu_1}{\sigma_1}\right)^2} + N_2 e^{-\frac{1}{2}\left(\frac{\Delta p_T - \mu_2}{\sigma_2}\right)^2} \right)$$

Each of the parameters N_2, m, and s are linear functions of the particle jet's transverse momentum. Since each of the five parameters needs two parameters to define its linear dependence, there are ten parameters in total. The best parameters are chosen by a likelihood fit over all the jets and applied as a correction.

One source of systematic uncertainty for this correction is the response of the electromagnetic calorimeter. This can be estimated by comparing data and Monte Carlo events for $W \to e\nu$ and $J/\psi \to ee$ events. The difference in the distribution of calorimeter energy

divided by the track momentum gives a systematic uncertainty of 3.5%. The detector simulation gets less accurate near the edges of calorimeter towers: calibration with $Z \to ee$ events shows a 10% discrepancy in these regions, which is taken as a systematic uncertainty. An additional uncertainty comes from the difference in particle multiplicity given by different showering algorithms; the difference between them adds a 1% uncertainty.

Out-of-cone and underlying event uncertainty

An additional source of systematic uncertainty comes from the fraction of a jet's energy that falls outside the cone of 0.4. This energy is modeled imperfectly in the Monte Carlo events, so a systematic uncertainty is assigned by examining photon + jet events in data and Monte Carlo. An annulus around the jet with a radius between 0.4 and 1.3 in the η–ϕ plane is examined, and the energy in this region is compared between data and Monte Carlo simulation. The largest difference between Monte Carlo events and data is taken as a systematic uncertainty. The uncertainty is derived as a function of transverse momentum because jets with higher transverse momentum are narrower; the uncertainty is as large as 4% for low-momentum jets.

Another source of systematic uncertainty is the contribution from what is called the *underlying event*: extra particles resulting from interactions between the remaining partons of the proton and antiproton. These effects are estimated by looking at tracks that form an angle between 60° and 120° with the jet. This region is sensitive to effects from radiation, multiple interactions, and beam remnants. Comparing this region in data and Monte Carlo

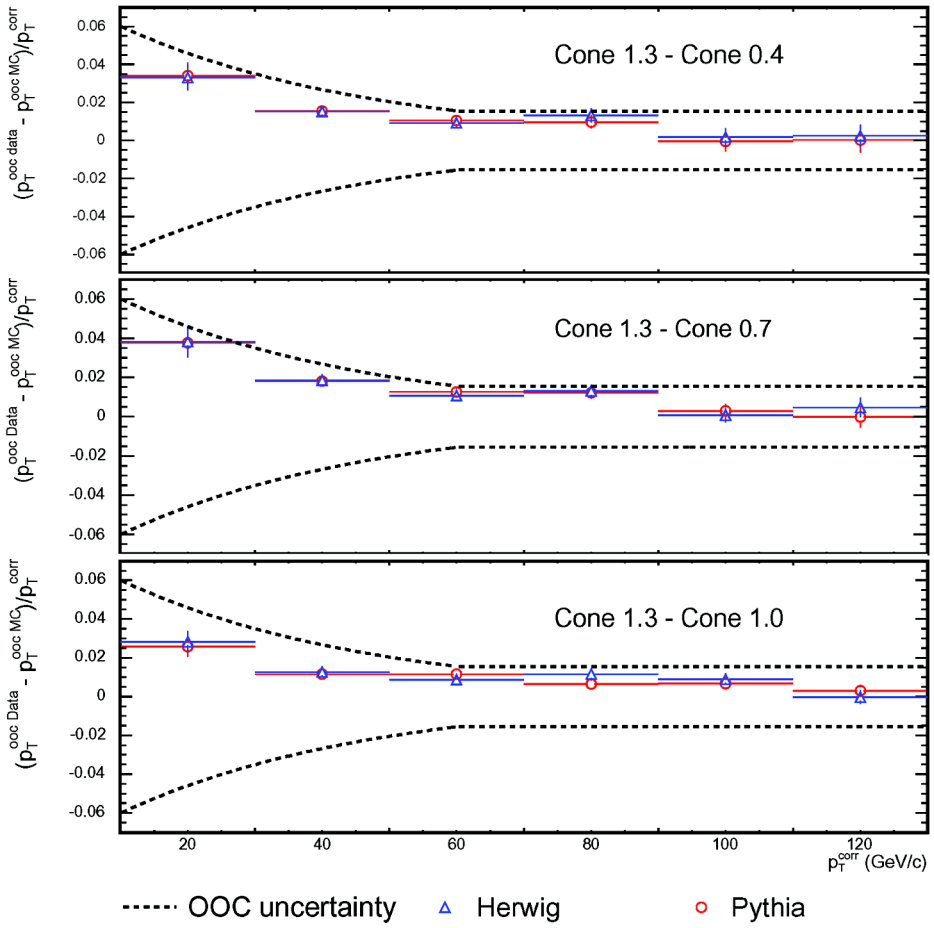

Figure 19. The uncertainty on the jet energy scale from out-of-cone energy, which covers the difference between data and Monte Carlo jets for all values of jet transverse momentum.

and taking the difference between them gives an uncertainty of 10% for low momentum jets, though the uncertainty decreases at higher energies.

Splash-out uncertainty

Finally, some of a jet's energy is still not included in the cone of 1.3 used for the out-of-cone correction. In Monte Carlo events, there is an average of 0.5 GeV of "splash-out"

energy that falls outside this region. Because measuring such small amounts of energy so widely spread out is very difficult in data, half of this value, 0.25 GeV, is taken as an additional systematic uncertainty.

Missing transverse energy

Neutrinos cannot be detected by CDF, but an energetic neutrino will manifest itself as missing energy. Physicists use a quantity called *missing transverse energy* by convention, although it might be more accurately characterized as missing transverse momentum. The transverse energy E_T of a given tower in the calorimeter is the energy in the tower times the transverse component of the vector that points from the center of the detector to the tower:

$$E_T = E \sin\phi \hat{n}$$

Here E is the energy in each tower, ϕ is the polar angle of the tower, and n is the unit vector directed from the center of the detector to the tower in the azimuthal plane. The missing transverse energy, usually written $\not\!\!E_T$, of an event is calculated with a vector sum over the transverse energies of the calorimeter towers: $\not\!\!E_T = -\sum_i E_{Ti}$, where i indexes each tower.

After the missing transverse energy is calculated, it needs to be corrected for the position of the primary vertex of the event, which, if not at the center of the detector, will cause an adjustment in the direction of the vector associated with each tower. The resulting two-dimensional vector gives a magnitude and a direction of a (massless) undetected particle. The longitudinal component cannot be determined by missing energy because the

longitudinal component of the momenta of the initial-state partons is not known, while the transverse component can be assumed to be nearly zero. Dealing with this requires some care in reconstructing the neutrino.

4.3 Luminosity calculation

The integrated luminosity of a given data period must be accurately measured to make a precise prediction of how many events of a given process are expected. The CLC luminosity monitor can be used to estimate the integrated luminosity, which can be calculated by the equation $R_{p\bar{p}} = \mu_{CLC} \cdot f_{BC} = \sigma_{inelastic} \cdot \varepsilon_{CLC} \cdot L$, where R is the collision rate, μ_{CLC} is the number of collisions per bunch crossing observed by the CLC, f_{BC} is the bunch crossing rate, $\sigma_{inelastic}$ is the inelastic proton-antiproton cross section, ε_{CLC} is the CLC acceptance, and L is the luminosity. The key term μ_{CLC} is determined from the number of bunch crossings in which the CLC sees no particles:

$$\mu_{CLC} = -\ln\left(\frac{N_{zeroBC}}{N_{totalBC}}\right)$$

The estimate for the cross section is based on previous measurements from CDF [64] and the acceptance is calculated based on data and simulation studies. Uncertainties in these two quantities dominate the total uncertainty on the luminosity, which is 6%.

Chapter 5

Event selection

Being able to measure the single top production cross section requires a good understanding of both the single top production signal and its backgrounds in order to distinguish one from the other. This requires a thorough analysis of the signal and the backgrounds which give rise to the same experimental signature. This section gives a general overview of the processes that contribute to the sample.

The final state of a single top production event has a W boson and two quarks, at least one of which is a bottom quark (two in the s-channel). The W boson can decay into either two quarks or a charged lepton and a neutrino. The branching ratio to quarks is twice as large as that to leptons; in addition, the tau lepton is very difficult to detect, making the practical branching ratio to leptons even smaller. (A tau lepton can be detected in this analysis when it decays to a muon or electron, about ten percent of the time.) However, leptons provide a very clean signature that makes it much easier to remove background events. This added sensitivity more than compensates for the smaller branching ratio, so this analysis looks only at leptonically decaying W bosons.

This gives a final state with a charged lepton, a neutrino, and two quarks, at least one of which is a bottom quark. Both s- and t-channel single top diagrams can easily radiate an

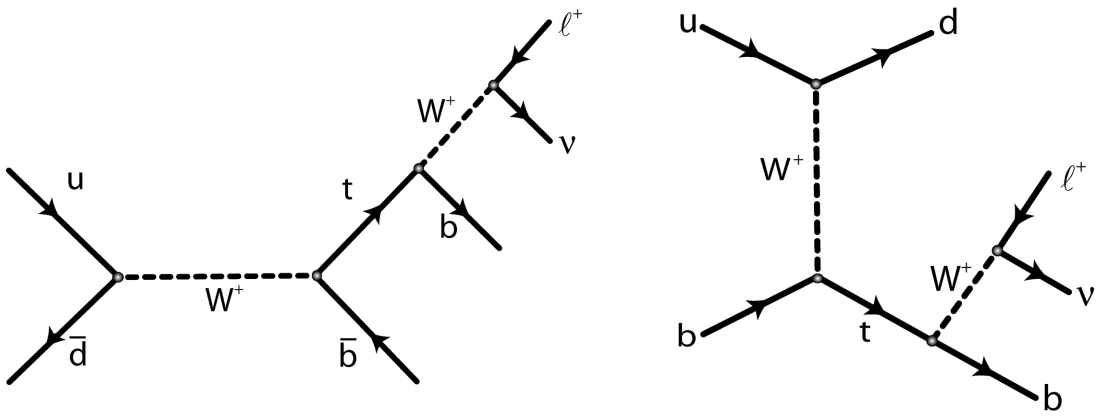

Figure 20. The *s*-channel (left) and *t*-channel (right) single top production diagrams used in this analysis..

extra gluon, so some events may have an additional gluon in the final state; thus, the desired event signature has two or three jets, at least one of which comes from a bottom quark; a charged lepton (excluding tau leptons); and a neutrino. The neutrino cannot be directly detected and thus manifests itself as missing transverse energy. Most of the background processes for single top also have a leptonically decaying *W* boson and *b* quarks in the final state.

A series of cuts must be applied to select events with this signature. This section details how these events are selected and how their efficiencies are modeled in Monte Carlo. This includes choosing a trigger, counting jets, selecting leptons, calculating missing transverse energy, rejecting backgrounds, and identifying jets which come from bottom quarks. However, selecting these events first requires an understanding of the background processes, so the cuts can be chosen to maximize signal and minimize background.

5.1 Major backgrounds

The major experimental challenge of a single top measurement is the large amount of background events that mimic the signal events. Top pair production, in contrast, has a clean signature of a single lepton, four energetic jets, and a neutrino that shows up as large missing transverse energy, and there is far more top pair production than anything else of events with this signature. Single top production, on the other hand, has a smaller cross section and only two jets, which makes it difficult to tell from the much larger backgrounds with the same signature.

W + jets

The largest background to single top production comes from quark interactions that radiate a *W* boson in association with two jets, which has the same final state as single top and a much larger cross section. Because most of the quarks in the final state are light quarks (up, down, or strange quarks), the level of this background can be reduced by *b*-tagging, a process which identifies jets with a displaced vertex that may have come from bottom quarks (see page 105). However, this fails to help in the case where a radiated gluon splits to two bottom quarks. This $W + bb$ background is the largest background process in this analysis. However, there is a substantial contribution from W + jets processes with lighter quarks in the final state; most of these are removed by *b*-tagging, but their cross section is very large, so they still form a substantial background.

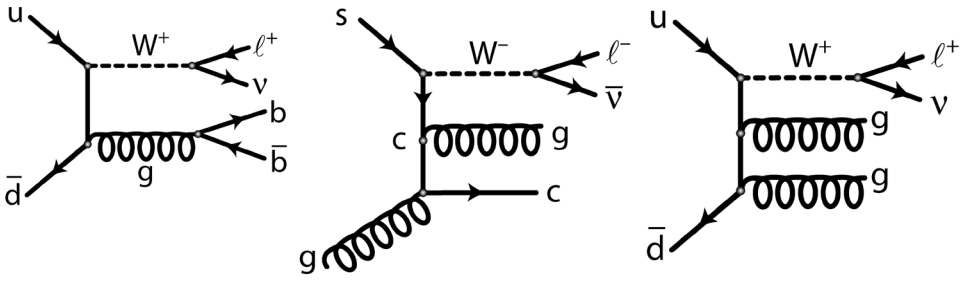

Figure 21. Three representative diagrams in the W + jets sample: (left) W + bb, (center) W + c + jet, (right) W + light flavor. These are the three major event topologies; there are many more diagrams in the W + jets sample.

Top pair production

Top pair production is a smaller background process, but nonetheless substantial. While no top pair decay has the same final state as single top, it can be misidentified if particles fail to be detected. This happens in dilepton top decays in which one of the leptons is missed, or in lepton + jets decays in which two jets are not detected. The resulting events, since they also come from top decays, are very similar kinematically to single top events, making this background difficult to deal with.

QCD multijet

Multijet events that contain no W boson have a very large inclusive cross section. In order to be confused with a single top event, one of the jets must "fake" an electron and a mismeasurement has to create a large missing transverse energy. While the probability of both these happening at once is very small, the large cross section of QCD multijet events makes this background non-trivial. Furthermore, because these events so rarely pass all

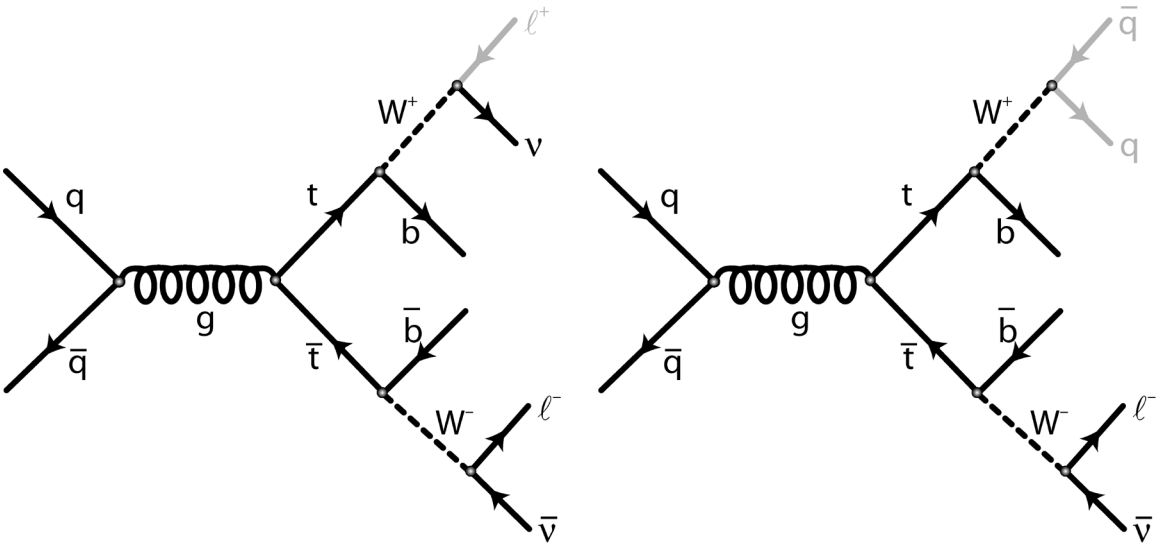

Figure 22. Feynman diagrams for top pair production, in (left) the case in which both *W* bosons decay to leptons and (right) in the case in which one *W* boson decays to quarks and the other decays to leptons. The grey particles indicates particles that must not be observed for the event to pass two-jet event selection.

selection cuts, it is not feasible to simulate these events with Monte Carlo, and its cross section is unknown. A data-based sample and estimate must be constructed.

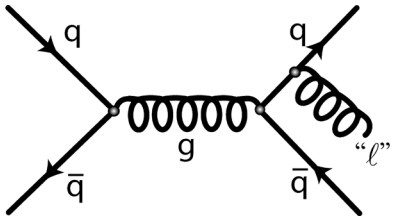

Figure 23. A sample diagram from the QCD multijet sample. To be accepted, a QCD event must have a jet from a quark or gluon pass all lepton selection requirements and be misidentified as a lepton.

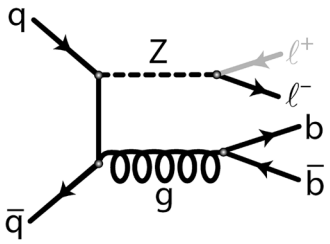

Figure 24. An example of a Z + jets diagram that is a background to single top production. One of the leptons from the Z boson decay must go undetected to pass the event selection requirements.

Z + jets

Since Z bosons decay to either two leptons or two neutrinos, Z + jets events do not often fake a single-top signature and the background is not large. However, because of its large inclusive cross section, some background remains from events in which a lepton is lost and its energy is counted as missing transverse energy. This background is considerably smaller, however, than the W + jets background.

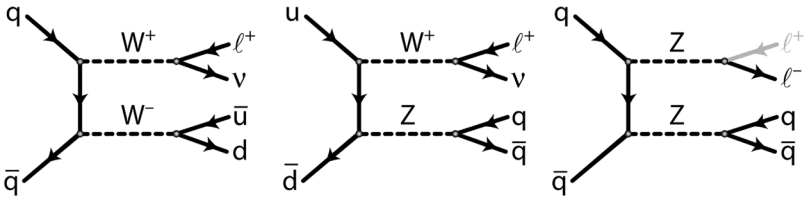

Figure 25. Three diagrams for diboson production: (left) WW production, (center) WZ production, (right) ZZ production.

Diboson

Electroweak diboson production, including WW, WZ, or ZZ production, also creates a small background, especially WW and WZ, which have the same final state as single top

Trigger	Integrated luminosity (pb^{-1})
Central electrons	2200 ± 130
Plug electrons	2200 ± 130
Central muons	2200 ± 130
Forward muons	2150 ± 130
Untriggered muons	2090 ± 125

Table 2. Integrated luminosity recorded for each trigger.

production. However, since their cross section is small, they constitute a small contribution to the total background.

5.2 Trigger

The first step of an analysis is to choose a trigger to examine. Data is written out by CDF along a variety of *trigger paths*, and choosing the right trigger is key to optimizing the event selection. Because the analysis has one charged lepton in the final state, it is sensible to trigger on leptons, because electrons and muons are the easiest particles to identify. This analysis uses two electron triggers, one for electrons detected in the central calorimeter and one for the plug calorimeter; two muon triggers, one for the CMU and CMP detectors and the other for the CMX detector; and a missing transverse energy trigger which increases the acceptance of muons. Each trigger is a composite of the three levels in the CDF trigger system (see page 39), with stricter requirements imposed at each level.

The triggers, while very efficient, will not trigger every event that they ought to. Thus, each trigger's efficiency must be estimated to properly predict the rate of each physics process. The basic approach to this is to examine a pure sample obtained through a different trigger, apply the trigger's selection cuts, and see how often the trigger for such events actually fired. The method for deriving of this efficiency is presented following the description of each trigger.

Because some triggers have prescales, and some have not been active for as long as others, the collected luminosity is different for each trigger. The luminosity is summarized in Table 2.

Central electrons

The central electron trigger (ELECTRON_CENTRAL_18) includes more data than any other trigger. The cuts used in the trigger for the three levels of trigger are summarized in Table 3. At Level 1, the extremely short time allowed to make a decision restricts the system to look for an energetic calorimeter cluster and an energetic track (not matched together).

- $\boldsymbol{E_T}$: The calorimeter towers are grouped into *trigger towers*, each 0.2 units of pseudorapidity in η and 15° in φ, to reduce the computational overhead of clustering. The transverse energy of the trigger tower must be high enough to pass this requirement.
- $\boldsymbol{E_{had} / E_{EM}}$: The ratio of energy deposited in the hadronic calorimeter to the energy

	Event variable	Selection requirement
Level 1	E_T	> 8 GeV
	E_{had} / E_{EM}	< 0.125
	p_T	> 8.34 GeV/c
	$N_{COT\ hits}$	≥ 4
Level 2	E_T	> 18 GeV
	$\|\eta\|$	< 1.317
Level 3	$Lshr$	< 0.4
	$\|\Delta z\|$	< 2.0 cm
	E_{had} / E_{EM}	< 0.125 for three towers
	p_T	9 GeV/c

Table 3. Event selection requirements for the central electron trigger.

in the electromagnetic calorimeter, which should be small for electrons.

- $\boldsymbol{p_T}$: The transverse momentum of the track.
- $\boldsymbol{N_{COT\ hits}}$: The number of hits made by the track in the COT tracking chamber.

At Level 2, the tower found at Level 1 is combined with nearby towers to create a cluster. The transverse energy of the cluster will naturally be larger than that of the single trigger tower. The track requirement remains the same.

- $\boldsymbol{E_T}$: The transverse energy of the cluster.
- $|\boldsymbol{\eta}|$: The pseudorapidity of the cluster. This requires the cluster to be in the central calorimeter.

	Event variable	Selection requirement
Level 1	E_T	> 8 GeV
	E_{had} / E_{EM}	< 0.0625
	Missing E_T	> 15 GeV
Level 2	E_T	> 20 GeV
	E_{had} / E_{EM}	< 0.125
	$\|\eta\|$	> 1.1 and < 3.6

Table 4. Event selection requirements for the plug electron trigger. The Level 3 requirements are the same as Level 2, but with fully reconstructed clusters and tracks.

Level 3 performs full reconstruction of the clusters and tracks.

- ***Lshr***: Short for "lateral energy sharing." This variable quantifies the difference between the pseudorapidity distribution of the energy in the calorimeter and what is expected for an electromagnetic shower. The expectation is derived from simulations and modified to fit test beam data. This requirement helps remove hadronic showers that might imitate electromagnetic showers.
- $|\Delta z|$: The difference in the z direction between the calorimeter cluster and the extrapolated track.

The efficiency of this trigger is estimated by examining events in a pure sample of W bosons which decay to an electron and a neutrino, selected from a trigger that uses a single electron and large missing transverse energy. The fraction of these events which also

fired the ELECTRON_CENTRAL_18 trigger is the trigger efficiency. The average efficiency in this dataset is 96.5 ± 0.4%.

Events chosen through the central electron trigger are often referred to as CEM events because they are matched to clusters in the central electromagnetic calorimeter.

Plug electrons

The plug electron trigger (MET_PEM) does not trigger directly on clusters in the plug calorimeter. This is because the plug has higher background from "beam splash"—particles from elastic collisions and collision remnants—which tend to have large pseudorapidity. In addition, the trigger cannot use tracking information because most tracks that point toward the plug do not pass through enough layers of the COT to make fast tracking feasible. Instead, this trigger relies on the presence of large missing transverse energy in the final state to trigger events. The lack of a tracking requirement for this trigger makes it less pure, thus requiring additional selection cuts to purify the sample. The trigger requirements are summarized in Table 4. Level 3 has the same requirements as Level 2, but with fully reconstructed data.

Two samples are used to estimate the efficiency of this trigger. To measure the efficiency of the calorimeter energy requirement, a sample of Z bosons that decay to electrons is used, with one electron triggered in the central electron trigger and the other detected in the plug calorimeter. Requiring that these electrons come from a Z boson makes this sample very pure. The average efficiency of this trigger is 88.6 ± 0.6%.

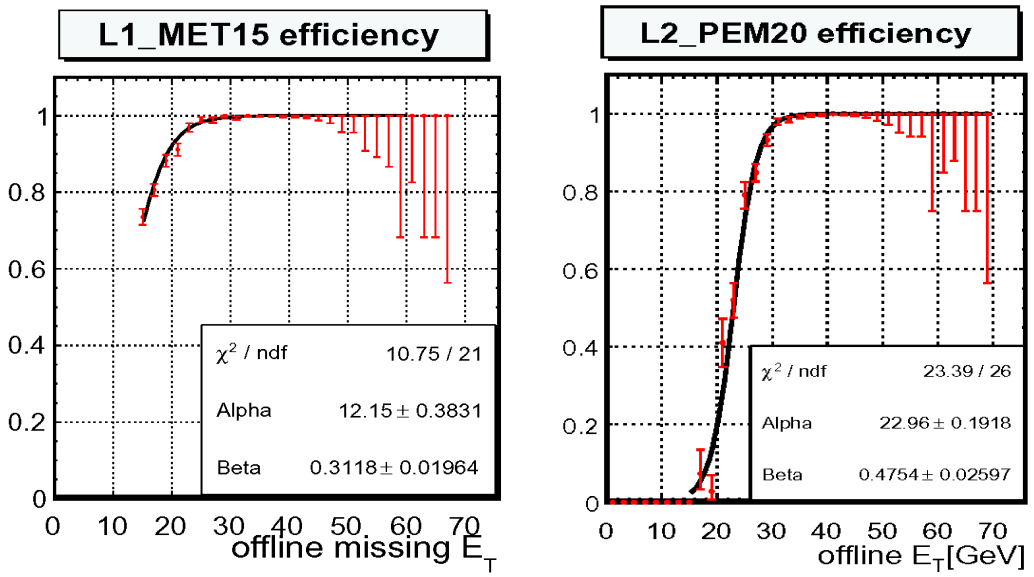

Figure 26. Turn-on curves as fit to data for (left) missing transverse energy and (right) electron transverse energy. These curves are applied to Monte Carlo events to make their kinematics match the data.

Missing transverse energy is calculated more carefully when processed offline than in the trigger, so the final value may lie below the value that the trigger uses. Because of his, Monte Carlo events must use a *turn-on* function to sculpt their kinematics to match the data. This function is parameterized in a *W* sample triggered with only a calorimeter energy requirement, modeled by the function $\varepsilon(x)=\frac{1}{1+e^{-\beta(x-\alpha)}}$. This parameterization gives a weight that is applied to Monte Carlo events in this region.

Because additional corrections are applied to plug energy measurements in offline processing which are not included in the Level 2 trigger calculation, the distribution of transverse energy in this sample also needs a turn-on curve. The curve is derived in the same way as the missing transverse energy curve and corrects the electron energy in this sample

	Event variable	Selection requirement
Level 1	$p_{T\,CMU}$	> 6 GeV/c
	CMP	At least one stub
	p_T	> 4.09 GeV/c
Level 2	E	Minimum ionizing energy
	p_T	> 14.77 GeV/c
Level 3	$\|\Delta x_{CMU}\|$	< 10 cm
	$\|\Delta x_{CMP}\|$	< 20 cm
	p_T	< 18 GeV/c

Table 5. Event selection requirements for the central muon trigger.

to match the data.

Because these events use the Phoenix tracking algorithm offline, they are often referred to as Phoenix electrons, or PHX for short.

Central muons

The central muon trigger (MUON_CMUP18) detects muons that reach the central muon chambers. There must be hits in both the CMU and the CMP to pass this trigger. The trigger requirements are summarized in Table 5.

- $\boldsymbol{p_{T\,CMU}}$: The transverse momentum of the stub in the CMU.
- $\boldsymbol{E}$: The total energy deposited in the calorimeter by the muon candidate.
- $\boldsymbol{p_T}$: The transverse momentum of the track.
- $|\boldsymbol{\Delta x}|$: The distance between the muon stub and the extrapolated track in the direction

perpendicular to both the beamline and the radial vector.

The efficiencies of the muon triggers are measured in events with Z bosons that decay to two muons, requiring the muons to have opposite charges and an invariant mass near the Z mass. With these events, one muon activates the trigger and the other one is examined to see if it also activated the trigger. Much information can be gained by examining the events in which one muon is detected in the CMU and CMP and the other is detected in the CMX. This comparison allows a simultaneous extraction of efficiencies for this trigger and for the CMX trigger. The efficiency for central muons is 91.6 ± 0.5 %.

Because these events must be recognized by both the CMU and CMP subdetectors, they are referred to as CMUP events.

Forward muons

The trigger for muons in the more forward CMX subsystem (MUON_CMX18_DPS) requires a muon stub matched to a COT track. Because it only has one set of detectors (instead of two, like the CMUP trigger) and no steel between it and the calorimeter, the CMX trigger is not as pure as the CMUP trigger. This gives it a relatively high rate of non-muon particles that are triggered. Because its overall rate is high, the CMX trigger must be prescaled in order to keep the trigger bandwidth reasonable. The selection requirements are summarized in Table 6. The variables used have been introduced already except for one:

- CSX: The scintillators around the CMX must give a timing signal consistent with particles coming from Tevatron collisions to reduce acceptance of muons from

	Event variable	Selection requirement
Level 1	$p_{T\,CMX}$	> 6 GeV/c
	p_T	> 4.09 GeV/c
	$N_{COT\,hits}$	≥ 4
	CSX	Pass timing requirement
Level 2	p_T	> 14.77 GeV/c
Level 3	$\|\Delta x_{CMX}\|$	< 10 cm
	p_T	< 18 GeV/c

Table 6. Event selection requirements for the forward muon trigger.

cosmic rays.

The efficiency of this trigger is measured at the same time as the central muons as already described. Its efficiency is 95.7 ± 0.6%.

Events recorded by this trigger are usually called CMX events. The integrated luminosity for this trigger is smaller than that of the other triggers because the CMX trigger was not included near the beginning of Run II, since it still needed to be studied and understood. In addition, the triggers are often prescaled, which affects the collected luminosity.

Untriggered muons

Muons tend to have fewer fake events than electrons because they are detected farther from the interaction point, have more material around them to absorb non-muons, and

	Event variable	Selection requirement
Level 1	Missing E_T	> 15 GeV
	Tower E threshold	> 1 GeV
Level 2	Seed E	> 3 GeV
	Tower E threshold	> 1 GeV
	E_T	> 10 GeV
	N_{jets}	≥ 2
Level 3	Missing E_T	> 35 GeV

Table 7. Event selection requirements for the missing transverse energy trigger.

require minimum ionizing energy in the calorimeter. Thus, there are many events in the muon systems that cannot be triggered efficiently but contain many muons that can be recovered offline with additional selection cuts. Adding these events requires the use of a trigger that does not use muons.

Without a lepton to use for the trigger trigger, the easiest way to identify events is to look for their other significant features: jets and missing transverse energy. One trigger(MET35_&_TWO_JETS) requires two jets and 35 GeV of missing transverse energy. The selection requirements are summarized in Table 7.

- **Missing E_T**: The missing transverse energy of the event. The requirement is lower

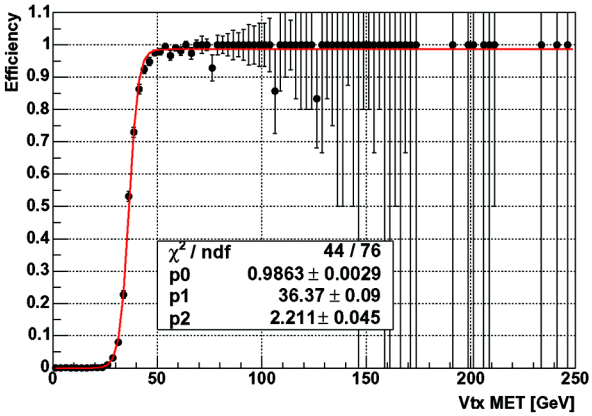

Figure 27. The turn-on curve applied to the untriggered muons as a function of missing transverse energy without muon corrections (Vtx MET).

for Level 1 because other higher-level triggers use the same Level 1 trigger.

- **Tower *E* threshold**: The energy needed in a tower to be included in the missing transverse energy calculation (Level 1) or the clustering calculation (Level 2).
- **Seed *E***: The energy of the highest-energy tower in a calorimeter cluster.
- $\boldsymbol{E_T}$: The transverse energy of a cluster in the calorimeter.
- $\boldsymbol{N_{jets}}$: The number of calorimeter clusters passing all cuts.

As instantaneous luminosity at the Tevatron increased, it was necessary to modify this trigger to require that one of the jets be central (MET35_&_CJET_&_JET), requiring $|\eta| < 1.1$. This, combined with a dynamic prescale, lowered the rate sufficiently to keep using the trigger.

In order to make this trigger fully efficient, additional kinematic cuts are required. Two jets with transverse energy greater than 25 GeV, corrected with jet energy corrections,

are required; one of the jets must be in the central region, with a $|\eta| < 0.9$; and the distance between the jets in the η–ϕ plane must be more than 1.0.

Because the offline missing transverse energy calculation includes energy from the muon, it can different dramatically from the calculation used in the trigger. Thus, using this trigger requires the Monte Carlo simulation to be adjusted by a turn-on function in missing transverse energy, similar to that used for the plug electron trigger. The turn-on is measured in events triggered by the CMUP muon trigger, comparing the number that pass a cut on offline missing transverse energy with the number that actually passed this trigger.

Events recorded by this trigger are often referred to as "loose muons," although *untriggered muons* is a more appropriate name. This trigger adds signal acceptance equal to about 40% of the CMUP trigger.

This trigger has less integrated luminosity than the other triggers because of the prescales applied to this trigger.

5.3 Jet multiplicity selection

Once an event is selected by a trigger, it must be reprocessed offline. The offline selection identifies jets in an event. The distribution of jet multiplicity of single top production events before any selection cuts is shown in Figure 28. This analysis examines events with two or three jets, thus selecting most single top events. Events with only one jet

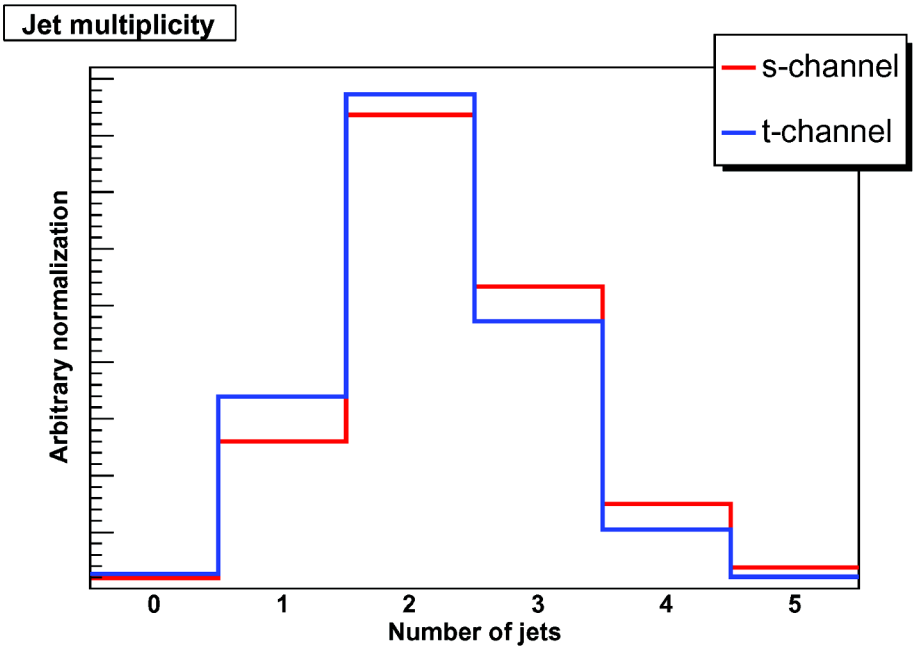

Figure 28. The number of reconstructed jets in Monte Carlo single top events. This analysis looks at events with two or three jets.

have a very large background from W + jets processes that make them of little use for measuring a single top production signal.

5.4 Lepton selection

The basic lepton selection cuts made by the trigger still leave a large number of fake leptons. Additionally, some triggers have no lepton requirement at all. Good lepton identification is vital to purify the sample by removing fake leptons, making it easier to understand and estimate the background to the single top signal.

One important variable for lepton identification is *isolation*. This quantity allows discrimination against leptons which form inside jets and do not originate from hard

scattering events. Isolation is defined by computing the transverse energy in a cone of radius 0.4 in the η–ϕ plane. The isolation is the ratio of the transverse energy that is not in the lepton cluster to the transverse energy in the cluster. A small number indicates that there is little extra activity in the calorimeter near the lepton, so it is unlikely to come from a jet. If this quantity is less than 0.1, the lepton is said to be isolated or *tight*; otherwise, it is non-isolated or *loose*. All leptons in this analysis are required to be tight, to pass the isolation criterion; however, loose leptons are still used to remove dilepton events (see page 101).

While all of the cuts shown here remove mostly background, they will also cut out some true leptons. Thus, it is important to estimate the rate at which these events can be reconstructed. This can be estimated by processing Monte Carlo events; however, due to the imperfect modeling of the detector, the efficiency will not be perfectly modeled. This necessitates the addition of a Monte Carlo scale factor, which is the ratio of the data and Monte Carlo efficiencies. Applying this factor to the Monte Carlo acceptance gives a proper estimate of the efficiency of lepton identification.

Central electrons

The largest sample of candidate events comes from the central electron trigger. A lepton candidate is a cluster of energy in the central calorimeter matched to an extrapolated track from the central tracker. A summary of the selection cuts is given in Table 8. A description of the variables in the table follows.

- **Geometry**: The candidate must have a cluster in the central calorimeter and the

Event variable	Selection requirement
Geometry	Fiducial in CEM
E_T	> 20 GeV
p_T	> 10 GeV/c
E_{had} / E_{EM}	< 0.055 + 0.00045E
E / p	< 2 if p_T < 50 GeV/c
$Q \times \Delta x$	> –1.5 cm and < 3.0 cm
$\|\Delta z\|$	< 3.0 cm
χ^2_{strip}	< 10
$Lshr$	< 0.2
Conversion	Pass conversion requirement
COT track quality	Pass COT track requirements
Isolation	< 0.1

Table 8. Event selection requirements for central electrons.

cadidate must be in the fiducial region; that is, it must be in a region that can be triggered with good efficiency. A cluster near a crack in the calorimeter will not be able to be triggered easily, so it is removed from both data and Monte Carlo to avoid difficulties of modeling such a region.

- $\boldsymbol{E_T}$: The transverse energy of the cluster. The trigger requires only 18 GeV; requiring 20 GeV ensures that the trigger is fully efficient.
- $\boldsymbol{p_T}$: The transverse momentum of the associated track. This requirement removes

many low-momentum electrons that come from *Bremstrahhlung*.

- $\boldsymbol{E_{had} / E_{EM}}$: The ratio of hadronic to electromagnetic calorimeter energy. Electrons leave very little energy in the hadronic calorimeter, so this cut removes background from hadronic objects. The energy-dependent term raises the threshold for high-energy clusters because high-energy leptons will leave more energy in the hadronic calorimeter.
- $\boldsymbol{E / p}$: The ratio of the cluster's energy and the track's momentum. This ratio should be nearly one for a true electron, so this requirement removes many fake electrons. For sufficiently high-momentum tracks, fakes of this sort are unlikely, so the requirement is relaxed.
- $\mathbf{Q \times \Delta x}$: The quantity Δx is the signed difference in x between the track and the cluster when the track is extrapolated to the position of the shower max, where x is a local coordinate defined to be perpendicular to both the particle beam and the radial vector to that calorimeter tower. Q is the measured charge of the particle. The asymmetry in the requirement results from the trajectory of particles in the detector; if the sign of the charge and Δx are opposite, the particle traverses a larger part of the calorimeter in adjacent towers, which results in more radiation and a less precise final position. This means the cut must be looser in this case to preserve signal efficiency.
- $|\mathbf{\Delta z}|$: The absolute value of the difference in z position between the cluster and the

extrapolated track.

- χ^2_{strip}: A single charged particle leaves a Lorentz shower profile in the CES shower maximum detector. The measured profile is compared, strip by strip, to the predicted profile, derived from theoretical parameterizations and test beam studies. The χ^2 is calculated between the predicted and measured profiles.
- ***Lshr***: This variable was already defined (see page 75).
- **Conversion**: Photon *conversions* are an important background for electrons. A photon traveling through material can convert into an electron-positron pair, and the electron, though a true electron, is not meaningful to the analysis because it comes from a photon and not a hard scattering event. The conversion veto looks for an track with the opposite charge of the electron track that is separated from it by less than 2 mm in the $r-\phi$ plane at the point at which they are parallel. It also requires the cotangent of the polar angle between the two tracks to be less than 0.04. If such a track can be found, the electron is likely to come from a photon conversion and is not accepted.
- **COT track quality**: The tracks in the COT must be of high quality. Each track must have hits in at least five hits in each of three axial superlayers and two stereo superlayers. This ensures that the track is cleanly reconstructed.

The identification efficiency of these cuts is measured in data using a sample in which Z bosons decay to an electron and a positron which are detected in the central

calorimeter. A cut around the *Z* boson mass makes this sample very pure; counting leptons with the same sign gives an estimate of the background. One of the electrons required to pass the trigger and identification cuts, and the other one is examined to see if it also passed the identification cuts. The fraction of identified electrons is the data efficiency and is around 80%.

The same procedure is done in a Monte Carlo simulation of the same process to measure the Monte Carlo efficiency. Because these two are not identical, their ratio is taken as an efficiency scale factor. The scale factor, averaged over all run ranges, for central electrons is 97.9 ± 0.5 %.

Plug electrons

Plug electrons, because the leptons are not triggered and because the tracking is less reliable, must have a different set of cuts to improve their purity. This sample has more fake leptons than any of the other samples, even after the identification cuts. A candidate event is a cluster in the plug matched to a track reconstructed by the Phoenix algorithm. The selection requirements are summarized in Table 9 and discussed thereafter.

- **Geometry**: The cluster must lie in the plug calorimeter.
- p_T: The transverse momentum of the lepton. Although this trigger contains leptons with lower momenta, they have a large background from fake electrons and are difficult to simulate properly.
- $|\eta|$: The absolute value of the pseudorapidity of the cluster. Outside of this range,

Event variable	Selection requirement
Geometry	Fiducial in PEM
p_T	> 20 GeV
$\|\eta\|$	> 1.2 and < 2.0
E_{had} / E_{EM}	< 0.05
χ^2_{PEM}	< 10
E_5 / E_9	< 0.65 for both layers
$\Delta R_{PEM\text{-}PES}$	< 3 cm
$N_{silicon}$	≥ 3
Isolation	< 0.1

Table 9. Event selection requirements for plug electrons.

reconstruction is inefficient has a large background from elastic proton-antiproton collisions.

- $\boldsymbol{E_{had} / E_{EM}}$: The ratio of hadronic to electromagnetic energy. Electrons leave nearly all their energy in the electromagnetic calorimeter.
- $\boldsymbol{\chi^2_{PEM}}$: The shower profile in the PEM detector is compared to the predicted Lorentz distribution and a χ^2 is calculated.
- $\boldsymbol{E_5 / E_9}$: Clusters resulting from true electrons tend to have a narrow energy distribution. This can be seen in the PES shower maximum detector, where the energy of the five strips at the center of the cluster (E_5) is compared to the energy of all nine strips in the cluster (E_9) for both layers of the PES. For an electron, most of

the energy will be contained near the center of the cluster, and this ratio will be large.

- $\Delta R_{PEM\text{-}PES}$: The distance in the x–y plane between the position of the reconstructed cluster in the PEM calorimeter and the PES shower maximum detector. This removes a background from poorly reconstructed clusters that might otherwise fake an electron.
- $N_{silicon}$: The number of hits in the silicon detector of the associated track. This requirement improves the quality of the track and reduces the background from poorly reconstructed tracks.

Estimation of the identification efficiencies is done in a way similar to the central electrons, except one electron is triggered in the central region and the other is required to be in the plug. The efficiency is lower in the plug region because of higher backgrounds and less efficient tracking, falling to around 70% in data. The scale factor for these electrons, averaged over all run ranges, is 91.4 ± 1.4 %.

Muons

The large amount of material to absorb other particles and the requirement of minimum ionizing energy in the calorimeter make the muon sample very pure. A common set of identification cuts applies to all muons, with additional cuts are required for each subdetector.

The common muon cuts are presented in Table 10.

- p_T: The transverse momentum of the track.

Event variable	Selection requirement
p_T	>20 GeV/c
E_{EM}	2.0 + max(0, 0.0115(p – 100)) GeV
E_{had}	6.0 + max(0, 0.028(p – 100)) GeV
d_0	< 0.2 cm; < 0.02 cm if no silicon hits
COT tracks	Pass COT tracking requirements
χ^2_{track}	< 2.3
Isolation	< 0.1

Table 10. Event selection requirements for muons.

- $\boldsymbol{E_{EM}}$: The energy deposited in the electromagnetic calorimeter. This requirement removes particles, especially pions, with enough energy to "punch through" the calorimeter, since they will leave more energy in the calorimeter. The second term, dependent on the particle's momentum p, accounts for the natural rise in ionization energy that a true muon will leave if its momentum is large, in accordance with the Bethe-Bloch equation [65].
- $\boldsymbol{E_{had}}$: The energy desposited in the hadronic calorimeter. This removes hadrons with enough energy to punch through the calorimeter. The scaling is different from E_{EM} because of the different material and thickness of the hadronic calorimeter.
- $\boldsymbol{d_0}$: The *impact parameter* of the track. This quantity is the distance between the beamline and the position of the track's reconstructed vertex in the r–ϕ plane. This cut removes a background from in-flight decays of long-lived particles into pions or

kaons by requiring that the track point back to the beamline. If there are no hits in the silicon detector, which indicates a lower quality track, the cut must be tighter.

- **COT tracks**: Tracks for muons in all subdetectors except the BMU must have at least three axial and at least two stereo COT superlayers with at least five hits each. This ensures that a good track is reconstructed. The BMU, because it is so far forward that tracks do not pass through as much of the COT, cannot use as strict a tracking requirement, and instead requires that 60% of the COT wires along the track have hits in them.
- χ^2_{track}: The track is compared to the position of the hits in the COT tracking chamber and a χ^2 is calculated. This reduces the background from poorly reconstructed tracks, primarily from kaons that decay in flight.

Because there are several different muon subdetectors and several different triggers, there are eight different muon types used in this analysis. Each of them has slightly different cuts to account for the specific characteristics and geometry of the subdetector. The selection requirements are summarized in Table 11. Most muon types have a requirement on Δx, which is the distance between the stub and the extrapolated track in the direction perpendicular to the beamline and to the radial vector to the cluster.

- CMUP, the primary muon trigger, requires stubs in both the CMU and CMP muon chambers. These chambers are the most central and the redundancy of the two systems allows for a very pure sample with high efficiency. The stubs are required

Muon type	Selection requirement
CMUP	$\Delta x_{CMU} < 7$ cm, $\Delta x_{CMP} < 5$ cm
CMX	$\rho < 180$ cm, $\Delta x_{CMX} < 6$ cm
CMU	$\Delta x_{CMU} < 7$ cm
CMP	$\Delta x_{CMP} < 5$ cm
BMU	$\Delta x_{BMU} < 9$ cm
CMXNT	$\rho > 180$ cm, $\Delta x_{CMX} < 6$ cm
CMIO	$E_{EM} + E_{had} > 0.1$ GeV
SCMIO	$E_{EM} + E_{had} > 0.1$ GeV

Table 11. Specific event selection requirements for each muon type.

not to fall within 3 cm of the most forward part of the CMP detector because the detector is not considered trustworthy in that region. .

- CMX, the forward muon trigger, requires stubs in the CMX subdetector. However, there is a section of the CMX that is not used in the trigger, wherever the radius is less than 180 cm, which is too far forward to get good tracking and has too high a fake rate to use in the trigger. As with the CMP, a CMX stub is required to be 3 cm from the forward edge of the detector to ensure its quality.

The remainder of the muon types are not triggered but are added through the missing transverse energy trigger. Although they have different quality requirements, after this stage they are collected into a single muon category.

- CMU muons have stubs in the CMU detector but not the CMP. These often occur

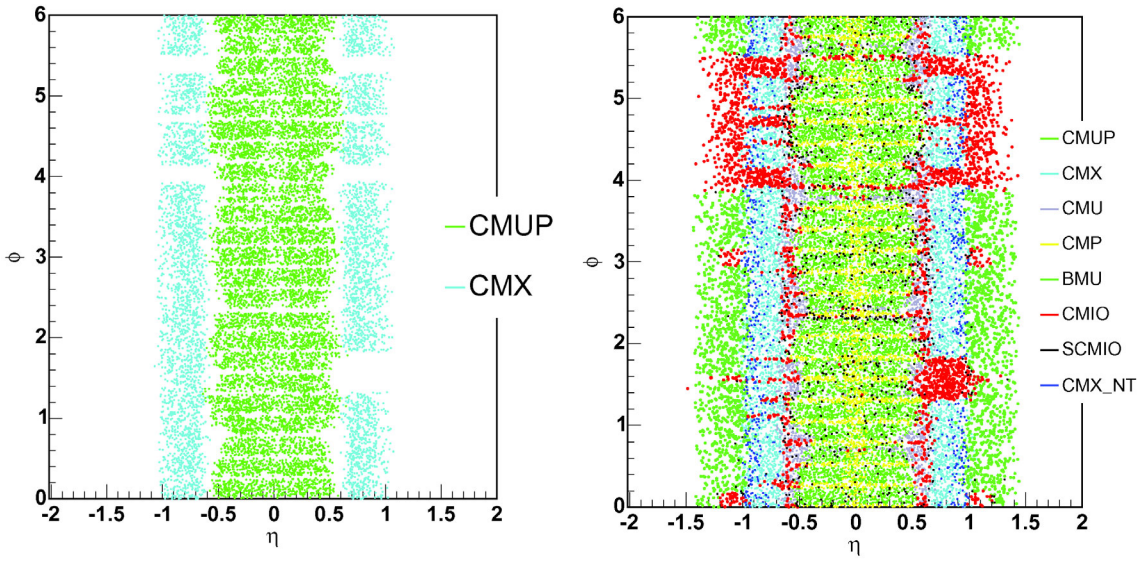

Figure 29. The distribution, in the η–φ plane, of muons from (left) CMUP and CMX triggers and (right) all muons, including untriggered muons added from the missing transverse energy trigger.

when muons land in the forward 3 cm of the CMP, which are rejected from the CMUP.

- CMP muons have stubs in the CMP detector but not the CMU. These show up because the CMP covers a broader range in pseudorapidity than the CMU. The fiduciality requirements are the same as for CMUP muons.
- BMU muons have hits in the BMU subdetector. Because it is the most forward of the muon detectors, the BMU's tracking is less precise, and it needs looser tracking cuts. The stub must be more than 3 cm from the forward edge of each chamber, and more than 13 cm if $|\eta| < 1.25$, which puts it in a less reliable region of the

subdetector. It is also required that Δx be less than 9 cm. This looser requirement results from the longer path length required by the muons to reach the detector.

- CMXNT muons leave stubs in the non-triggerable region of the CMX detector. They have the same quality requirements as the normal CMX sample.
- CMIO muons are isolated tracks matched to calorimeter clusters that do not point toward a muon detector. These stubs are still required to be matched to a track in the COT and a low-energy calorimeter cluster. An additional energy requirement strengthens the requirement of minimum ionizing energy in the calorimeter.
- SCMIO muons are similar to CMIO muons but are matched to a stub in a non-fiducial region of the detector. The same quality requirements apply to these muons as to CMIO muons.

The identification efficiency of CMUP and CMX muons is measured in a way similar to their trigger efficiency: Z decays to muons are examined, with one leg triggered and the other examined to see if it was identified. The efficiency in data is about 92% in both regions. This results in a scale factor of 92.6 ± 0.6 % for CMUP muons and 99.3 ± 0.7% for CMX muons (as well as CMXNT muons, which have no difference in efficiency).

The same method is used to calculate identification efficiencies for the untriggered muons, requiring one muon to be triggered in the CMUP. The resulting scale factors are 89.1 ± 1.4 % for the CMU, 92.0 ± 1.2% for the CMP, 112 ± 0.9% for the BMU, 104 ± 1.3% for the CMIO, and 98.6 ± 1.6% for SCMIO muons.

5.5 Missing transverse energy

Missing transverse energy, the signature of a neutrino, is an important part of event selection cuts because it removes events from many QCD multijet processes that have no final-state neutrino. Combined with the lepton identification requirement, a missing transverse energy cut makes it very likely that an event included a W boson in its final state that decayed leptonically. Because of this, it is important that missing transverse energy be calculated as accurately as possible.

At the trigger level, the missing transverse energy is corrected for the position of the reconstructed primary vertex. Another correction must be included for an event with a muon, since the muon carries substantial energy but leaves little in the calorimeter. Thus, the transverse energy of the muon must be included in the calculation. However, since a muon is also matched to a calorimeter cluster with minimum-ionizing energy, the energy of its associated cluster must be subtracted from its total energy, or else the calorimeter energy would be counted twice. Furthermore, when jet corrections are applied to jets, the missing transverse energy must be corrected as well.

This analysis requires the missing transverse energy to be greater than 25 GeV, which removes a large portion of the QCD multijet background. However, events that do not pass this selection requirement are useful for estimating the remaining QCD multijet contribution.

5.6 Event vetoes

Several additional selection requirements are required to remove specific backgrounds. Each of these vetoes removes a large portion of a specific background to improve the purity of the final sample. This analysis removes events with a primary vertex too far from the center of the detector. Additional vetoes take care of events from cosmic rays, *Z* bosons, processes with two leptons, and QCD multijet processes.

Primary vertex requirement

For all leptons, the *z* coordinate of the reconstructed vertex of the track must be less than 60 cm from the center of the detector. This requires the tracks to come from a hard scattering process and not elastic scattering or cosmic rays. This affects the luminosity calculation because some true hard scattering events occur outside this region. Measurements in data, using a trigger on events with hits in the CLC, show that the luminosity should be scaled by 96.4 ± 0.4 %.

Cosmic ray veto

Muons coming from decays of cosmic-ray pions in the upper atmosphere pass through the detector frequently. Because these are true muons, they pass the muon identification requirements; because they do not originate from collisions in the Tevatron, they need to be identified and removed. This requires a series of additional cuts.

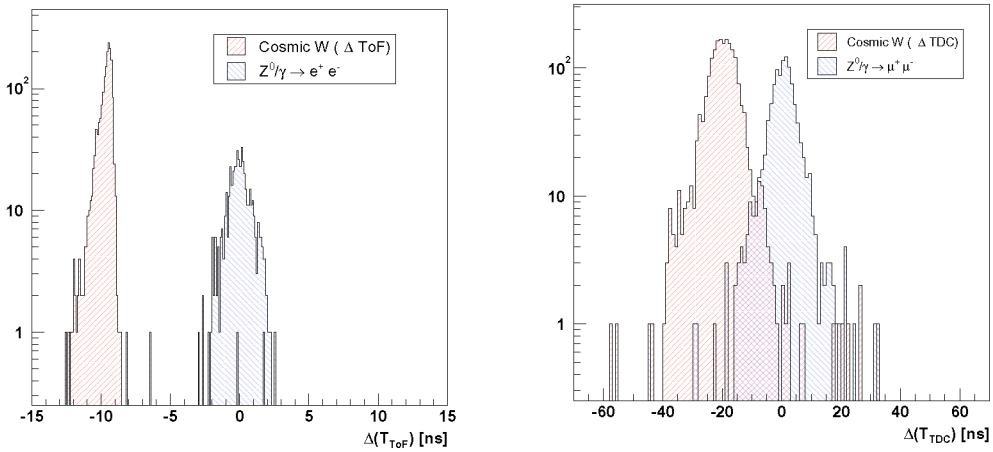

Figure 30. Variables used to remove cosmic ray events, shown for cosmic rays and true *Z* boson events: (left) the different in TOF measurements, and (right) the difference in hadronic calorimeter TDC measurements.

Because cosmic-ray muons pass through the detector from the top to the bottom, they often show up as back-to-back tracks. The impact parameters d_0 of the tracks will be essentially identical because they come from the same particle. This is most easily seen by converting to polar coordinates (ρ, β), where $\rho = \sqrt{d_{01}^2 + d_{02}^2}$ and $\tan\beta = \frac{d_{01}}{d_{02}}$. A cosmic ray will have impact parameters that are very similar but have a large angle between them. This analysis rejects events with $|\rho| < 0.2$ cm, $|\beta - \frac{3\pi}{4}| > 0.2$, and $|\beta - \frac{7\pi}{4}| > 0.2$.

Because the timing of a cosmic ray track with respect to bunch crossings is different from that of a track from a hard scattering event, often only one track will be reconstructed because the other fails the timing requirement for a COT track. This other track can be recovered by a dedicated algorithm which fits the COT hits and timing information under four different assumptions: the particle associated with the top or the bottom track can be

traveling either inward or outward. If the best fit is to the case of the top track heading inward and the bottom heading outward (the signature of a downward-going cosmic ray muon), and if the fit c^2 is less than 300, the event is removed.

Cosmic rays will also likely not pass near a reconstructed vertex. If the weighted mean of the z position of the vertices in an event is more than 4 cm from the average z_0 position of the tracks reconstructed by this method, the event is removed.

The TOF timing system can also distinguish cosmic rays by comparing the times of flight of the two tracks. If the bottom track's time of arrival is more than 5 ns after the top track, the event is identified as a cosmic ray event and removed.

The hadronic calorimeter TDC timing system is also used to identify cosmic-ray events, though its resolution is not as good as the TOF's. If the difference in time between the top and bottom clusters in an event is more than 20 ns, the event is removed. This cut is tightened to 10 ns if two muon stubs are detected and the difference in polar angle between them is more than 3 radians.

These cuts cause a reduction in true hard-scattering events of about 1.5 % in the muon sample, and leave behind at most 3% cosmic background contamination in the inclusive W sample, most of which is easily removed by other selection cuts. This veto is only applied to data because cosmic ray events do not appear in Monte Carlo events..

Z boson veto

Z bosons can form a significant background because sometimes, if a *Z* boson decays to two leptons, one lepton is lost, and mismeasured jets can result in sufficient missing transverse energy to pass the missing transverse energy requirement. The *Z* boson veto looks for a very loosely identified lepton (including a high-momentum track with no other tracks nearby) with the opposite charge of the identified tight lepton. If the invariant mass of the two falls near the *Z* boson mass (76–106 GeV), the event is rejected. This leaves very little residual contamination from Z + jets events.

Dilepton veto

One major background is top pair production in which both final-state *W* bosons decay to leptons. If one of the leptons ignored, the signature is the same as a single-top event. This background can be greatly reduced by removing any events with more than one lepton. This veto looks for any leptons, including loose leptons, in an event, and rejects the event if there is more than one lepton.

Because leptonically decaying top pair events have two true leptons in their final state, they require a special scale factor to account for the difference between data and simulation in the rate of misidentifying one lepton. This can be calculated by measuring the rates of identification of all possible pairs of leptons and calculating a combined dilepton

veto scale factor, which is 1.08 ± 0.02. This is applied only to top pair production Monte Carlo events with two leptonic decays.

QCD multijet veto

The safest way to deal with the difficult QCD multijet sample is to remove as much of it as possible. These events often have very strong kinematic features not described by other Monte Carlo samples that can be observed as a difference between data and Monte Carlo prediction. Because the missing transverse energy does not come from a neutrino, these mismatches often show up in angles between missing transverse energy and other objects.

One strong requirement for single top production is that the lepton and neutrino must come from a *W* boson decay. Thus, the *transverse mass*, the transverse component of the invariant mass, is a useful quantity, defined as

$$m_T = \sqrt{2\left(p_{Tl}p_{Tv} - p_{xl}p_{xv} - p_{yl}p_{yv}\right)}$$

(The invariant mass cannot be used because the neutrino's *z* momentum cannot be calculated.) This should have a peak at the mass of the *W*, about 80 GeV, for all true *W* events. Because the missing transverse energy in QCD multijet events comes from mismeasured jets, it often points the opposite direction as the fake lepton, which gives such events a low transverse mass. Applying a cut of 10 GeV on the transverse mass removes a large part of the QCD multijet background. This purifies the muon sample well (since it is

very hard to fake a muon stub and a calorimeter cluster with minimum ionizing energy), except for the SCMIO sample, which requires a cut of 20 GeV because of greater observed contamination.

The central electron sample, lacking the advantages of the muon detectors, has a more significant contamination from QCD multijet events. A transverse mass cut at 20 GeV removes much of the background, but some angular variables show some excess in the data . More discrimination can be provided by a variable called *MET significance*, defined as

$$METsig = \frac{\not\!\!E_T}{\sqrt{\vec{E}_{T\,unclustered} \cdot \hat{\not\!\!E}_T}}$$

The denominator is the amount of *unclustered* energy—energy not included in reconstructed jets—that points in the direction of the missing transverse energy, and acts as a measure of the uncertainty of the calculation. Other key variables are angles between missing transverse energy and jets or leptons. Comparing scatterplots of these variables in data and Monte Carlo events reveals disagreement in regions of small MET significance and transverse mass, which are expected to be rich in QCD multijet events. These regions can be removed by carefully chosen cuts (Figure 31). It is useful to define "triangle cuts" that cut a specific region in a two-dimensional space; this analysis uses two such cuts: it requires $METsig > 0.05 m_T + 3.5$ and $METsig > 2.5 - 3.125 \Delta\varphi_{MET,jet2}$. These cuts greatly reduce the amount of QCD multijet in the sample while retaining approximately 95% of the signal.

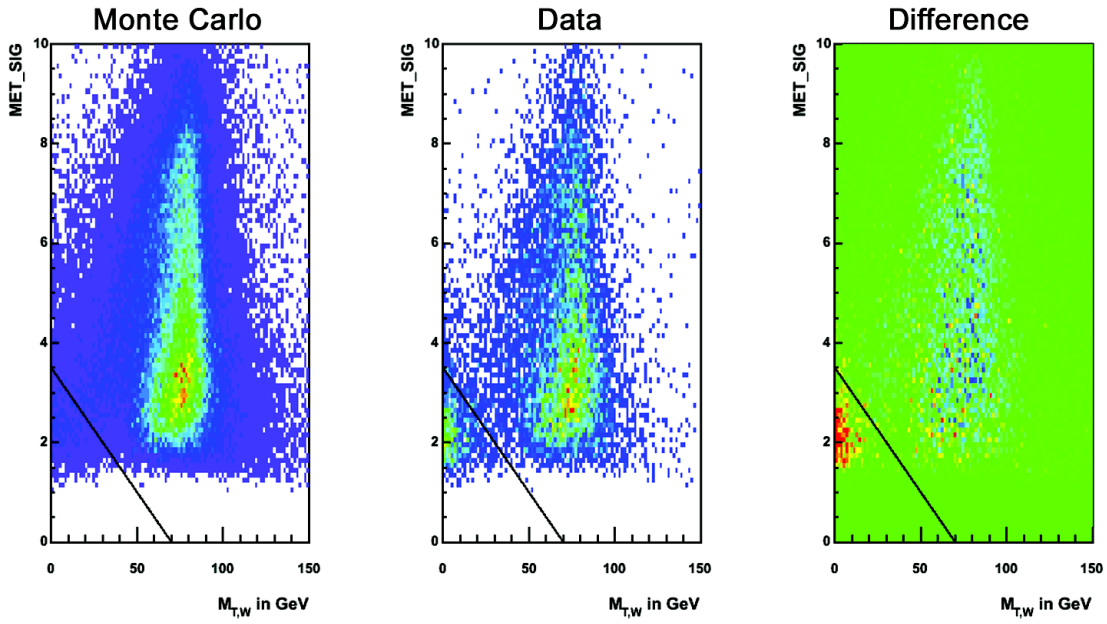

Figure 31. An example of the motivation for one of the QCD multijet veto requirements. The data (center) have an excess in a QCD-enriched region that the Monte Carlo events (left) do not have. This can be seen clearly by subtracting the two plots (right). The black line indicates the position of the selection requirement.

The Phoenix sample, due to the lower quality of its tracks, has more contamination from QCD events and requires a tighter cut. In addition to a transverse mass cut of 20 GeV, it requires that MET significance be larger than 2. It also requires $\not\!\!E_T > 45 - 30\Delta\phi_{MET,jet}$ for all jets in the event. This improves the modeling and dramatically reduces the estimated QCD component.

5.7 *b*-tagging

The vast majority of background events contain only light quarks in their final states. Determining if a jet is one that originated from a bottom quark (called a *b-jet*) is very useful for removing backgrounds. This procedure, called *b-tagging*, takes advantage of the fact that *B* hadrons can only decay through weak interactions and thus have a relatively long lifetime. Combined with the fact that these jets often have large transverse momentum and thus a large Lorentz boost relative to the lab frame, *B* hadrons travel an average of 6.8 mm transversely before they decay.

Displaced decays are a key part of recognizing jets from bottom quarks. Here the silicon detector demonstrates its usefulness, as it has a high enough resolution to detect tracks coming from a *secondary vertex*, slightly displaced from a primary vertex. While the efficiency of tagging *b*-jets is only about 40%, due to tracking resolution effects, the efficiency of tagging light jets is less than 1%, so the sample is highly purified.

The tagging algorithm, called SECVTX, uses a complicated set of selection requirements that are summarized in Table 12. First, it examines tracks coming from the primary event vertex and places a series of track quality requirements on these tracks:

- **# SVX *r–φ* hits**: The number of hits in the layers of the SVX silicon detector in which the strips are parallel to the beamline (*r–φ* layers). This gives the necessary resolution from the silicon detector to attempt *b*-tagging.

	Selection requirement	Tight SECVTX tag		Loose SECVTX tag	
		Pass 1	**Pass 2**	**Pass 1**	**Pass 2**
Track quality	# SVX r–φ hits	≥3	≥3	≥2	≥3
	χ^2 / d.o.f.	< 8.0	< 8.0	< 8.0	< 8.0
	p_T (GeV/c)	> 0.5	> 1.0	> 0.5	> 1.0
Displaced track selection	d_0 (cm)	> 0.15	> 0.15	> 0.15	> 0.15
	Δz_0 (cm)	> 2.0	> 2.0	> 2.0	> 2.0
	d_0 significance	> 2.0	> 3.5	> 2.0	> 3.0
Vertex creation	Seed vertex χ^2	< 50	—	< 50	—
	Attachment significance	< 4.0	—	< 6.0	—
	Track pruning χ^2	< 45	< 30	< 90	< 1000
Vertex quality	Highest track p_T (GeV/c)	> 1.0	> 1.5	> 1.0	> 1.5
	Vertex fit χ^2	< 50	< 50	< 120	< 2000
	L_{xy} significance	> 9.0	> 9.0	> 6.0	> 6.0

Table 12. The cuts used for the tight and loose SECVTX b-tag algorithms.

- **χ^2 / d.o.f.**: The χ^2 of the track's hits compared to its predicted trajectory, divided by the number of degrees of freedom. This requirement ensures that the track was reconstructed well.
- **p_T**: The transverse momentum of the track. Higher-momentum tracks often have better resolution.

These tracks are matched to reconstructed jets by associating each track with the nearest jet in η–φ space. Then the tracks associated with a given jet are put through displaced track selection requirements:

- $\boldsymbol{d_0}$: The impact parameter of the track, defined earlier (see page 92). Here a large value indicates a displaced track.
- $\boldsymbol{\Delta z_0}$: The difference in the z direction between the primary vertex and the origin of the track, another sign of a displaced track.
- $\boldsymbol{d_0}$ **significance**: The impact parameter divided by its uncertainty. This gives a tighter requirement on well-measured tracks.

If there are at least two such tracks that pass all requirements, the jet is said to be *taggable*. Taggability is a useful requirement in constructing control regions. Before the tagging algorithm continues, any pair whose invariant mass is consistent with a K^0 or Λ particle is removed. These long-lived particles can also lead to displaced vertices.

Having selected candidates for a displaced vertex, SECVTX begins its first pass at reconstructing a secondary vertex. Pairs of tracks are combined to form a secondary vertex if they meet the **seed vertex** $\boldsymbol{\chi^2}$ requirement, which is the χ^2 comparing the best-fit position of the secondary vertex with the vertices of the two tracks. Other tracks are attached if they pass the **attachment significance** requirement, which is their impact parameter significance with respect to the new vertex. Tracks are removed if their addition increases the χ^2 of the vertex fit by more than the **track pruning** $\boldsymbol{\chi^2}$ requirement.

Once a secondary vertex candidate is formed, it must pass a series of vertex quality requirements:

- **Highest track p_T** : The highest transverse momentum of a track associated with the vertex. This prevents poorly measured tracks from falsely being identified as coming from a secondary vertex.
- **Vertex fit χ^2**: The χ^2 comparing the best-fit position of the secondary vertex with the vertices of the tracks associated with it. This ensures that the tracks are consistent with originating from the secondary vertex.
- **L_{xy} significance**: The transverse displacement from the primary vertex, L_{xy}, divided by its uncertainty. This determines whether the vertex is sufficiently far from the primary vertex to be tagged as a secondary vertex.

If the new vertex passes all these requirements, the jet is marked as *b*-tagged. If the jet fails to be tagged in this pass, SECVTX attempts a second pass, collecting all the tracks in the jet and attempting to form a vertex with all of them. The selection requirements are now slightly different (as shown in Table 11); if the vertex can pass the same vertex quality requirements, the jet is marked as *b*-tagged.

The requirements used in this analysis are referred to as *tight b*-tagging cuts. A *loose* version of this algorithm uses the same method with less stringent requirements. This increases acceptance of *b*-jets but also increases the rate of tagged light jets. This analysis

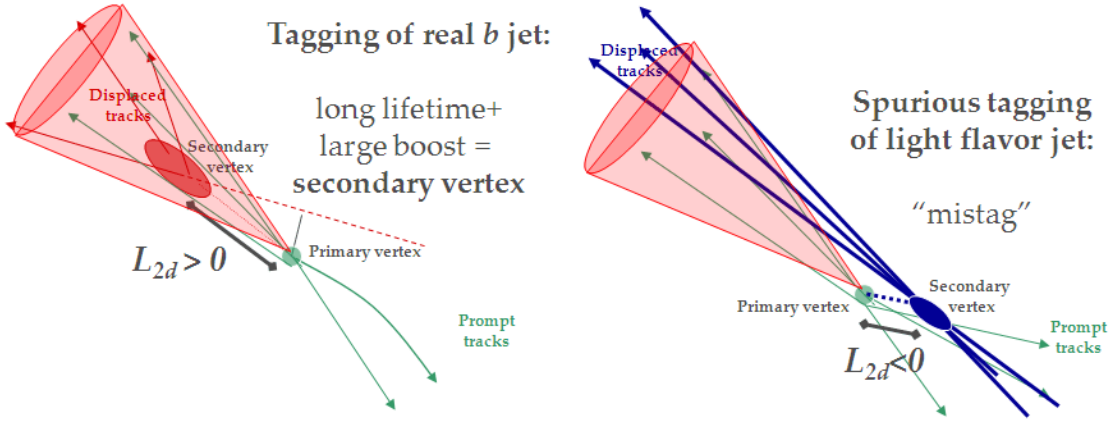

Figure 32. A diagram of a secondary vertex tag. A true *b*-jet (left) has a positive value of L_{xy}; a fake *b*-jet (right) has a negative value.

requires tight *b*-tags; however, loose tags are useful in studies when more tagged events are needed.

The quantity L_{xy} is useful when given a sign. It is defined to be positive when the secondary vertex is displaced in the same direction as the jet. A positive value of L_{xy} is consistent with a decay of a long-lived particle which results in a jet; a negative value of L_{xy} indicates an improper assignment of a *b*-tag dues to improperly reconstructed tracks. These *negative tags* are useful for estimating the rate of false *b*-tags.

Scale factors

Unfortunately, the extreme complexity of this technique means that the detector simulation is not able to simulate all details of *b*-tagging—simulated Monte Carlo events typically have more tagged *b*-jets than the actual data. However, this overestimated efficiency seems to have no kinematic dependence; it affects the rate of tagging but not the

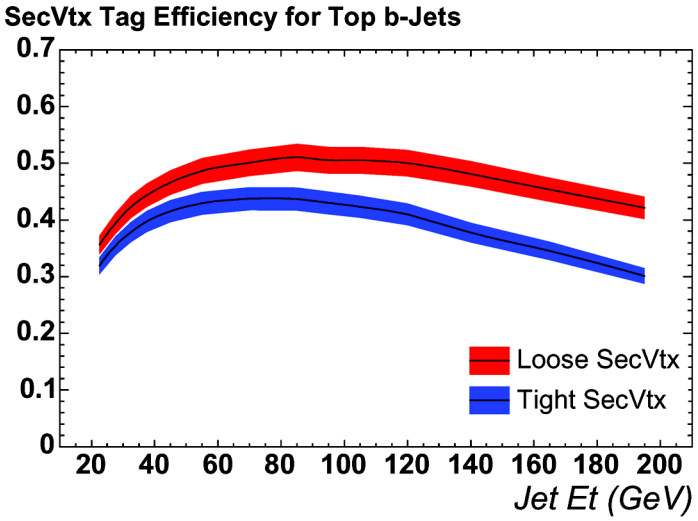

Figure 33. Tagging efficiency as a function of the transverse energy of the jets, for both the loose and tight versions of the SECVTX *b*-tagger.

kinematic distributions of tagged events. This means that it can be compensated for with a scale factor on the tagging efficiency.

This scale factor is estimated in two different ways, one using electrons and one using muons. The electron method uses a sample of two jets, each with transverse energy larger than 15 GeV. One of them (the *away jet*) is required to be tagged by the secondary vertex tagger; the other one (the *electron jet*) is required to contain an electron with transverse momentum of at least 9 GeV. The high-momentum electron in the electron jet makes it likely that it comes from a semileptonic decay of a *b* quark; requiring the away jet to be tagged increases the purity of the sample, since *b* quarks often come in pairs and double-tagged events are very rarely faked. The tagging rate of the electron jet allows an estimate of the tagging efficiency.

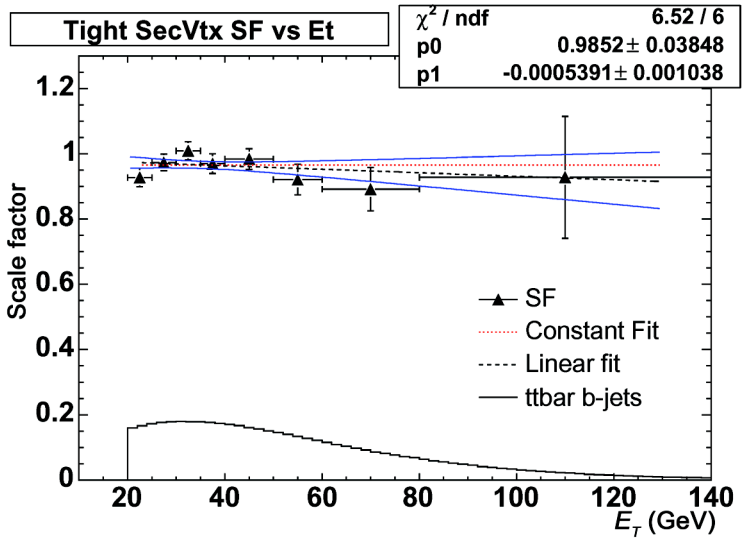

Figure 34. The *b*-tagging scale factor shown as a function of jet transverse energy. The ratio of the data and Monte Carlo events is consistent with a flat line.

The electron jets can be compensated for mis-tagged light quarks by looking for electron conversions in such jets; extrapolating this to the away jets gives an estimate of their heavy flavor fraction. The efficiency is then $\frac{N_+ - \alpha N_-}{N f_{hf}}$, where N_+ and N_- are the number of positive and negative tags, respectively, in the electron jets; N is the total number of events; f_{hf} is the heavy flavor fraction in the away jets; and α is a mistag asymmetry factor which must be derived in other studies (see page 114). The ratio of this quantity in data and Monte Carlo is the scale factor.

The muon method works on the same basic principle as the electron method, requiring an 8-GeV muon inside one jet and requiring the other jet to be tagged. The heavy flavor fraction is determined by performing a fit to the transverse momentum distribution,

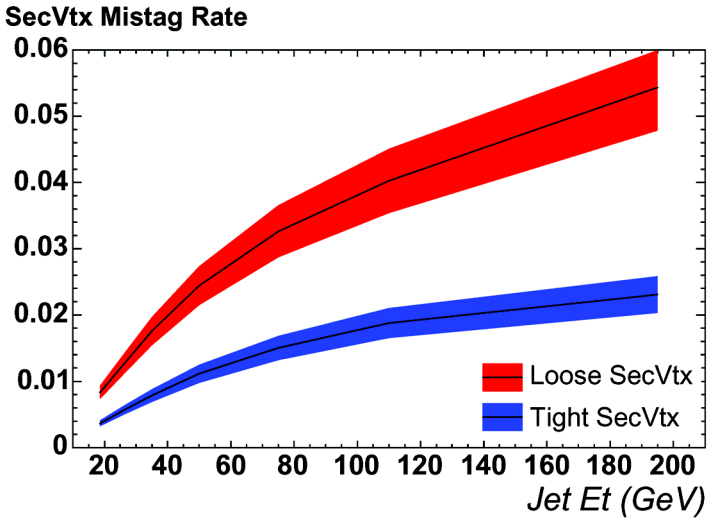

Figure 35. The rate of mistags for the loose and tight SECVTX *b*-taggers as a function of transverse energy. To keep the rate low, this analysis uses the tight SECVTX *b*-tagger.

using templates derived from Monte Carlo samples. This method gives consistent results with the electron method, and the two are combined to produce the overall scale factor of 95 ± 4 %.

The same scale factor is assumed for the tagging rate of charm quarks, which are tagged less frequently than bottom quarks but more so than light quarks. However, because this sample cannot be easily compared to data, the systematic uncertainty on the scale factor is twice as large.

Mistags

An important part of *b*-tagging is the accidental tagging of jets that do not contain true bottom quarks, called *mistags*. Even though the rate of mistags is very low, there are

far more events with light jets than with *b*-jets, causing a substantial contamination of the tagged sample. Unfortunately, the detector simulation does not properly model the rate of this contribution, and unlike the true tags, the mistag rate mismodeling has a strong dependence on kinematic variables.

To properly understand mistags requires a *mistag matrix*, which is a parameterization of the mistag rate as a function of several variables: transverse energy, the number of tracks in the jet, the sum of the transverse energies of all jets in the event, pseudorapidity, the number of reconstructed vertices in the event, and the *z* position of the primary vertex. Each variable is divided into four to eleven bins and used to construct a matrix of the rate of negative tags as a function of these six variables. The numbers in the matrix are calculated in a sample of generic jets.

The negative tag rate, however, is not the true rate of mistags. Some negative tags results from true jets whose tracks were badly reconstructed. The mistag matrix gives the negative tag probability measured in inclusive jet data, which is

$$R^{negTag} = \frac{N_{light}^{negTag} + N_{heavy}^{negTag}}{N_{light}^{total} + N_{heavy}^{total}}$$

The desired mistag rate is $\frac{N_{light}^{tag}}{N_{light}^{total}}$. Therefore, correction terms are needed. These are called *asymmetry* terms because they account for the difference between positive and negative tags. They are parameterized by two correction factors

$$\alpha = \frac{N_{light}^{tag}}{N_{light}^{negTag} + N_{heavy}^{negTag}} \text{ and } \beta = \frac{N_{light}^{total} + N_{heavy}^{total}}{N_{light}^{total}}.$$

Physically, α corrects for the presence of negative-tagged *b*-jets, while β corrects for the presence of *b*-jets in the generic jet sample used to derive the matrix.

The parameter α is derived from a likelihood fit to the data of the invariant mass of the tracks resulting from the displaced vertex. This variable shows good separation between *b*, *c*, and light jets. Templates are generated for each type of jet and then fit to the measured distribution. To remove some difficulties with low statistics, the distribution used in the fit is the *tag excess*, the positive tag distribution minus the negative tags. After this fit, all negative tag templates are scaled to match the observed rate, resulting in a negative tag scale factor, which is assumed to be the same for all jet flavors. After this scale factor is applied, the fit is performed again using both negative and positive tags, and this fit is used to calculate the parameter α. This parameter is also found to have a dependence on the jet's transverse energy, so it is calculated in four different regions of transverse energy.

The parameter β is derived from the same fit in the positive tag region, adjusted by the tagging efficiency and the Monte Carlo scale factor. The Monte Carlo scale factor is applied to the *b* and *c* samples equally.

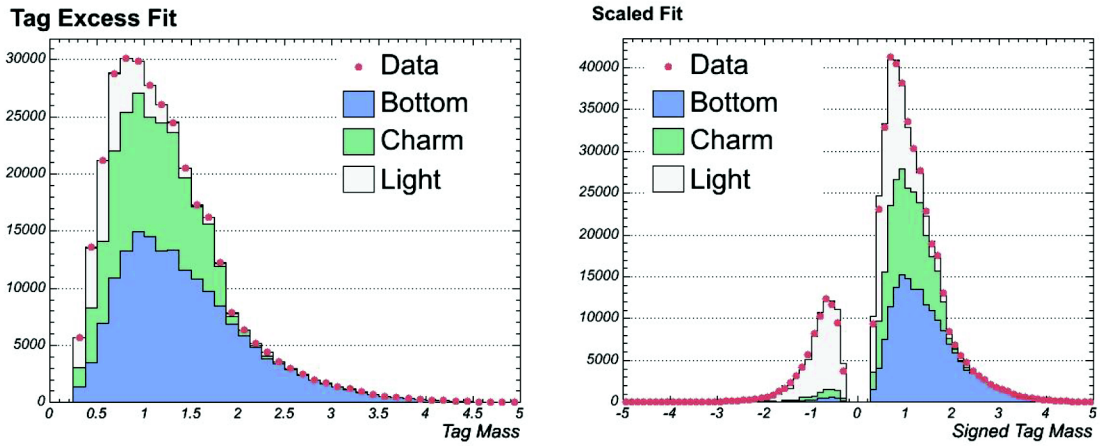

Figure 36. Plots used to derive the mistag asymmetry factors α and β: (left) the tag excess fit and (right) the scaled fit, both given as a function of the invariant mass of the secondary vertex.

An identical process is applied to templates made in a variable called *pseudo-ct*, defined as $L_{xy}\frac{m}{p_T}$, where all terms refer to properties of the secondary vertex. This similar to the proper decay length, but it only looks in two dimensions (hence the prefix "pseudo"). The difference between the result obtained from this variable and the invariant mass variable is taken as a systematic uncertainty.

Chapter 6

Modeling of processes

Once the event selection is finalized, each process that makes up the signal and background must be carefully modeled in Monte Carlo simulation to make sure it matches the data as well as possible. Large samples of simulated data are created for each signal and background process. Each sample is then passed through the same event selection to estimate the expected event yield and predict the kinematic distributions for each process.

6.1 *s*-channel single top

Single top Monte Carlo is generated by MadEvent [66]. This generator preserves information from the polarization of the top quark, which PYTHIA does not include. Studies by Sullivan [67] show that next-to-leading-order corrections change the cross section for *s*-channel but do not change any of its underlying kinematic distributions, so *s*-channel events are generated at leading order and scaled to the next-to-leading-order cross section.

6.2 *t*-channel single top

The same studies show that *t*-channel events change some of their kinematic distributions substantially with next-to-leading-order corrections. This is because the leading-order diagram has a *b* quark in the initial state (called a $2\to2$ process, since there are

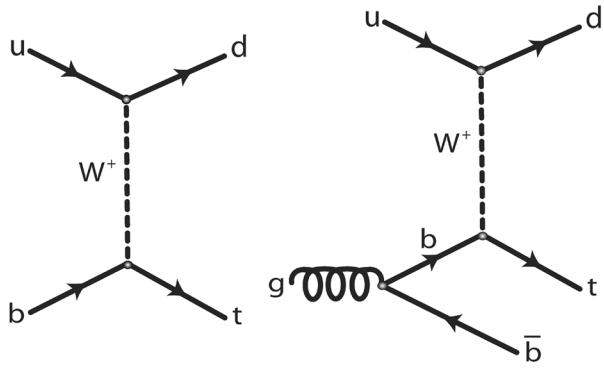

Figure 37. (left) The leading-order 2→2 *t*-channel diagram and (right) the next-to-leading-order 2→3, *t*-channel diagram.

two particles in the final state) whereas one next-to-leading-order diagram starts with a gluon that splits to a pair of *b* quarks (called a 2→3 process, since there are three particles in the final state) (Figure 37). These two contributions predict markedly different distributions of transverse momentum of the lower-momentum *b* quark.

To solve this problem, the *t*-channel sample is simulated by generating Monte Carlo events for both 2→2 and 2→3 processes using MadEvent. At generator level (before any showering or hadronization), the transverse momentum distributions of the lower-momentum bottom quark of these two samples are compared to a full next-to-leading-order distribution generated by the program ZTOP [68]. Adjusting their relative contribution reveals a point at which the two distributions intersect at 20 GeV. The *t*-channel sample is constructed by taking Monte Carlo events from the two samples in the fitted ratio, only using 2→2 events below 20 GeV and only using 2→3 events above 20 GeV. This gives a continuous

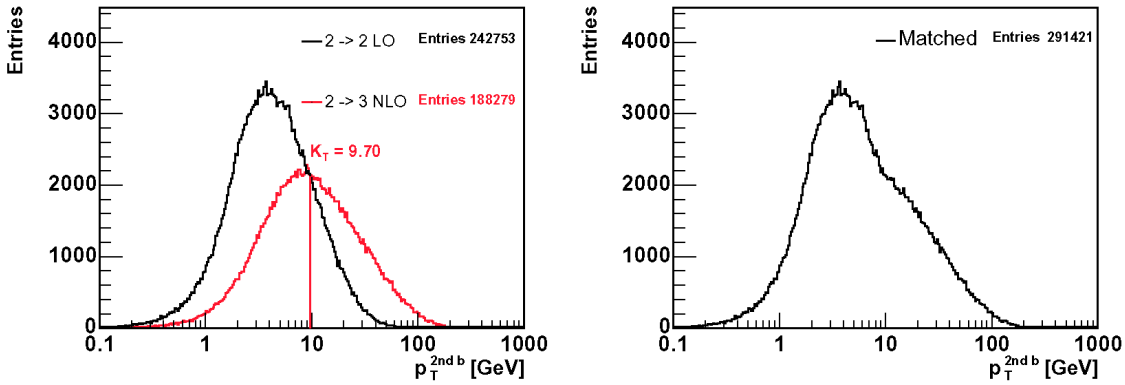

Figure 38. Illustration of the matching procedure for the *t*-channel Monte Carlo sample. (left) The 2→2 and 2→3 samples matched at the point at which they overlap. (right) The final distribution of the transverse momentum of the second quark.

distribution in the transverse momentum that matches the theoretically predicted distribution [69].

6.3 Top pair production and diboson samples

The top pair production background and the *WW*, *WZ*, and *ZZ* processes are generated by PYTHIA and normalized to the next-to-leading-order cross section. These samples are modeled well in PYTHIA.

6.4 *W* + jets

The largest background to deal with comes from *W* + jets production. Even at tree level (with no loops or renormalization calculations required), this background is described by millions of possible Feynman diagrams which describe the color and kinematic characteristics of the radiated gluons. ALPGEN is used to generate these events because it

properly calculates all tree-level matrix elements with full color and spin correlation information, which it passes to the PYTHIA showering routine.

Parton-jet matching

The PYTHIA showering algorithm gives rise to a difficulty when used with ALPGEN because there is an overlap in their generation of events. ALPGEN generates events at the matrix element level with initial- and final-state radiation, while PYTHIA approximates the effects of radiation by its showering. PYTHIA performs much more showering than ALPGEN does, but the initial stages of showering overlaps: ALPGEN might produce a diagram with a W boson and two radiated gluons, or it could produce a diagram with a W boson and one radiated gluon, while PYTHIA adds another gluon through parton showering. Because both these cases can occur, these events will appear with too large a rate.

The solution to this is a method usually referred to as *MLM matching* (for its inventor, Michelangelo L. Mangano). In this method, after parton showering, the final-state particles are grouped into jets by a jet-cone clustering algorithm, which groups all particles within a certain region of η–ϕ space. Each jet is then *matched* to a parton—a jet and a parton are associated if the parton lies within the cone of the jet. Only one parton can be matched to each jet. An event is rejected if it cannot match every parton to a jet.

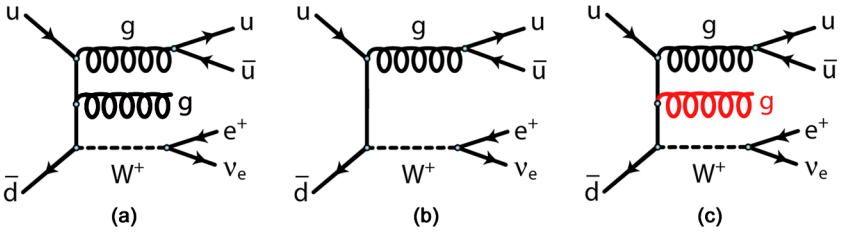

Figure 39. An illustration of the double-counting problem caused by using ALPGEN with PYTHIA. ALPGEN produces events from diagrams (a) and (b). PYTHIA's showering routine will sometimes takes events from diagram (b) and add a radiated gluon to produce diagram (c). In effect, this causes the diagram to be generated twice.

To get the counting right, the number of jets is required to be the same as the number of partons. Exclusive samples are generated with different numbers of jets, then added together after matching is performed. This removes double-counting in the showering.

Heavy flavor overlap removal

There is another problem of double-counting when using PYTHIA with ALPGEN, and it involves quarks with non-trivial masses, usually referred to as *heavy flavor*: charm and bottom quarks. It is important to separate events with these quarks because their kinematic behavior is different from the lighter quarks. However, they can arise in two different ways: they can be created at the matrix-element level in a $W + bb$ event, or they can arise from gluon splitting in the parton shower from a W + light flavor event. Because there is no difference between these two cases (they have the same Feynman diagram), combining ALPGEN and PYTHIA will overestimate the heavy flavor rate by counting the same events in both $W + bb$ and W + light flavor samples.

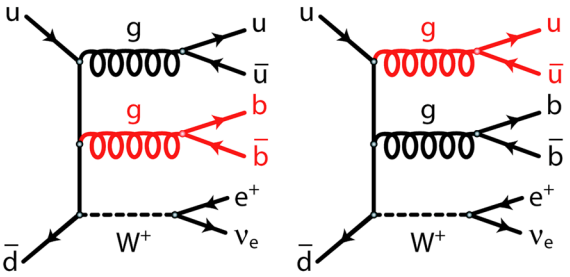

Figure 40. An example of the problem of heavy flavor overlap. The diagram on the left is generated as a W + light quark event to which PYTHIA adds a bottom quark pair during parton showering. The diagram on the right is generated as a $W + bb$ event to which PYTHIA adds a light-quark pair during parton showering. Since these cases result in the same diagram, the events will be double-counted.

The scheme for removing the overlap of heavy flavor divides heavy flavor events into two disjoint sets based on matching to fully reconstructed jets. Heavy flavor events generated by the matrix element are kept only if the heavy quarks lie in two different jets, while events generated by the parton shower are kept only if the heavy quarks lie in the same jet. This division is motivated by the expectation that quarks from showering will usually be close to their parents, while quarks from the matrix element are more likely to be well separated. This prescription removes the overlap between these events.

Heavy flavor separation

The W + jets sample consists of events generated from four processes: W + light quarks, $W + bb$, $W + cc$, and $W + c$. Each process is generated with up to four extra partons and the cross section is calculated for each process. These events are then combined into a single sample, weighting the events in each process by its relative contribution to the total cross section, to produce the final W + jets sample.

One consequence of this matching is that it is not possible to distinguish between an event with a W + light jet that radiates a gluon that decays into bottom quarks and a $W + bb$ event that radiates a gluon that decays into light quarks, because they represent the same Feynman diagram (as in Figure 40). However, separating the light quarks from the bottom quarks is useful, since $W + bb$ is a much more significant background process than W + light jets. Since it is not possible to split the sample based on the underlying process from which the events came, the W + bottom sample is defined as any W + jets event in which one jet falls within 0.4 in η–ϕ space of a bottom quark. Any remaining event which has a charm quark that falls this distance from a jet is classified as a W + charm event, and all remaining events are classified as W + light flavor.

Mistags

Modeling the contribution of W + light jet events that are nonetheless b-tagged—the mistag sample—is difficult because while the Monte Carlo simulation does a good job of predicting general W + jet kinematic shapes, some events are more kinematically disposed to mistagging, and the Monte Carlo does not model their kinematic features well. Therefore, the W + light flavor sample is not tagged directly; rather, the tagging requirement is relaxed to require only one taggable jet, and each event is weighted by the product of the mistag probabilities of its taggable jets. This results in a kinematic distribution that closely models the kinematic shapes of the mistag sample.

Because events with one and two tagged jets are treated separately, it is necessary to have a different mistag model in each region. The single-tagged model is constructed as described above; the double-tagged sample uses only events with two taggable jets, both weighted by their mistag probability.

6.5 Z + jets

The Z + jets process has the same difficulties as W + jets and is generated by ALPGEN in the same way. However, since its contribution to this analysis is small (because one of the leptonic decay products of the Z boson must remain undetected and generate sufficient missing transverse energy), it is not separated by quark flavors as the W + jets sample is, since the extra discrimination is not needed.

6.6 QCD multijet

The most difficult events to model come from QCD multijet events. The extremely high cross section of QCD multijet events means that even kinematically unlikely configurations can form a significant background. In particular, the conspiracy of improbable events needed to fake the necessary signature—a three-jet event in which one jet manages to pass all lepton cuts and, simultaneously, the energies are so badly measured that a large missing transverse energy is reported—still occurs enough to contaminate the sample significantly, even after the QCD veto. Because of the extremely small probability of these events occurring, and because they come from a hodgepodge of different QCD

processes, all of them difficult to calculate or model, it is impossible to simulate these events with Monte Carlo events.

Because these events cannot be simulated, the models are derived from data samples. This is complicated as well, because different lepton types may have vastly different rates and shapes of QCD multijet events, and each needs to be examined separately. This analysis uses three different models for QCD multijet events. All of them are based on the principle that QCD multijet events must contain a jet that is falsely identified as a lepton. Thus, by looking at jets that are not leptons but come close to passing electron cuts, it is possible to create a model of this background.

Jet-electrons

One strategy for creating a model of this sample uses a sample of generic jets, triggered through a generic jet trigger which simply looks for clusters of energy in the calorimeter. Since QCD multijet events must involve a jet that is falsely identified as an electron, this sample is examined for jets that look similar to electrons. Specifically, they must be energetic, having a transverse energy of at least 20 GeV; they must have a high fraction of energy deposited in the electromagnetic calorimeter, between 80% and 95%; and they must have fewer than four tracks, since a true electron has only a single track.

This jet is then assumed to be an electron and all the other event selection cuts are applied. Because jets have multiple tracks, their charge cannot be determined, so the electron charge is assigned randomly. Because they comes from generic jets similar to

electrons, these events are referred to as *jet-electrons*. Although they do not fit all kinematic variables well, after the QCD veto is applied they show a good agreement with data. The greatest drawback of this sample is its small size; only a few hundred events pass all the cuts required besides the *b*-tagging requirement, and only a handful are tagged as *b*-jets. However, unlike the other samples, it is possible to obtain a sample for the forward electron region, which the other methods cannot reach.

Anti-electrons

Another strategy for studying this sample uses the same trigger as the central electron sample. This strategy identifies electron cuts which depend on the kinematic properties of the event, such as transverse momentum, and others which rely only on detector effects, such as the fraction of energy deposited in the hadronic calorimeter. Five such "non-kinematic" cuts are identified: E_{had} / E_{EM}, χ^2_{strip}, $Lshr$, $Q \times \Delta x$, and $|\Delta z|$. These variables are designed primarily to reject fake electrons but do not greatly affect the kinematic properties of an event. Thus a QCD multijet model is constructed of events which fail at least two of the non-kinematic cuts but pass all kinematic cuts. This fake electron is chosen as the candidate electron, and the rest of the event selection cuts are applied.

Because these events are similar to electrons with some selection cuts inverted, they are given the unfortunate misnomer *anti-electrons*. Their advantage is their good match with kinematic variables, especially missing transverse energy and the angle between it and

observed particles. The biggest drawback of this method is its small size; as with the jet-electrons, only a few hundred events are collected.

Anti-electrons and jet-electrons give very similar kinematic shapes after the QCD veto is applied. Because of this similarity, and because the size of the samples is similar, both samples are used, added together to increase sensitivity. Because anti-electrons use the central electron trigger, only jet-electrons can be used to describe forward electrons. For muons, anti- and jet-electrons do a remarkable job of modeling the kinematic properties of the QCD multijet sample, and so the same events are used for the muon sample, with a cut on the events' pseudorapidity to confine them to regions appropriate to a given subdetector.

Non-isolated events

While the anti-electrons and jet-electrons do a surprisingly good job of modeling the QCD multijet contribution to the missing transverse energy spectrum for triggered muons, they do a poorer job in the untriggered muon sample, since these muons are kinematically sculpted by their trigger. For this sample a better model is obtained by using *non-isolated* events, events which pass all selection criteria except the requirement of lepton isolation. This is based on the rationale that non-isolated events are typically leptons contained in jets, and jets that contain energetic leptons are more likely to pass lepton identification cuts. This sample has the advantage of a large size; it is not used for the other triggers because it does not properly model key kinematic distributions, such as missing transverse energy.

However, in the untriggered muon sample it does a good job, probably because of the trigger requirements of jet separation and high missing transverse energy.

b-tagging

The statistics of the QCD multijet samples are too small for direct tagging; only a handful of jet-electron and anti-electrons events contain a tagged jet, making them useless for modeling kinematic distributions. The *b*-tagged model for the QCD multijet sample is estimated simply by using the distribution requiring at least one taggable jet. This matches the missing transverse energy distribution in the data of the QCD-enriched sample of low missing transverse energy. In addition, the measured tagging rate in data shows no significant dependence on missing transverse energy, indicating that there is no large change in the kinematic distributions of this sample after tagging.

Chapter 7

Predicted event yield

To be able to extract a single top signal requires a thorough understanding of the composition of the sample of candidate events. Both signal and background processes need to be carefully estimated in order to determine the sensitivity of the analysis is and the significance of an observed signal. Properly estimating each component of the background is essential to making a measurement of the signal top cross section.

The single top sample is broken up into events with exactly two or three jets, and with one or more *b*-tags. This gives a total of four categories. Each category has a separate estimate of the sample composition. Treating each sample separately improves the sensitivity of the analysis by combining regions with different signal purities. In addition, backgrounds are calculated for the orthogonal *untagged* sample, which requires at least one taggable jet, but no tagged jets; one-jet events, an important control sample of the W + jets background; and four-jet events, which are useful for validating the top pair production background.

7.1 Monte-Carlo-based background estimation

For any process for which the theoretical cross section is well-known, estimating the expected number of events in the sample is straightforward: $N = \sigma\varepsilon L$, for a cross section σ, efficiency ε, and integrated luminosity L, where the luminosity has been scaled by the primary vertex position scale factor. The efficiency is the fraction of generated events that pass all selection cuts, which can be calculated by generating Monte Carlo events and counting the fraction of events that pass the selection cuts. This number needs to be multiplied by the scale factors described previously: trigger efficiency, lepton identification scale factor, and *b*-tagging scale factor.

Calculating the *b*-tagging rate requires special care. Each jet is assigned a weight based on whether it can be matched to a heavy-flavor hadron, which means that the reconstructed jet lies within 0.4 in the η–ϕ plane of a hadron containing a heavy quark (bottom or charm) before detector simulation. If a jet is matched to a heavy flavor hadron and tagged, it is given a weight equal to the *b*-tag scale factor. If it is matched to heavy flavor but not tagged, it is not included in the sample and given a weight of zero. If the jet is not matched to heavy flavor, it is assigned a weight equal to its mistag probability, regardless of whether or not it was tagged, because the Monte Carlo simulation does not properly model mistagging. Untaggable jets always have a weight of zero.

Process	Cross section (pb)
s-channel	0.88 ± 0.05
t-channel	1.98 ± 0.08
Top pair	6.7 ± 0.83
WW	12.4 ± 0.25
WZ	3.96 ± 0.06
ZZ	1.58 ± 0.02
Z + jets	787.4 ± 50.0

Table 13. Cross sections used in this analysis for Monte-Carlo-derived backgrounds.

Each event is given a tagging probability. This probability depends on how many tags are assumed. The tagging probability is the jet weight w if the jet is assumed to be tagged, and $1 - w$ if the jet is assumed to be untagged. In some cases, such as the single-tag assumption in two-jet events, multiple combinations of jets must be examined, in this case: $w_1(1-w_2)+(1-w_1)w_2$. All possible combinations for a given tagging assumption are calculated for each event. Thus, each Monte Carlo event has a probability of no tags, exactly one tag, or two or more tags. This probability is used as an event weight in the drawing of histograms and in the calculation of the background estimate.

This method allows the calculation of a background estimate for all models with a well-understood cross section. This includes the signal *s*- and *t*-channel processes and the top pair production, diboson, and Z + jets processes. Single-top and top-pair cross sections

are calculated using a theoretical next-to-leading order calculation [70], assuming a cross section of 175 GeV/c^2. Next-to-leading order cross sections are also used for the diboson samples [71], and the CDF inclusive Z + jets cross section measurement [72] is used as the cross section for the Z + jets background. The results are summarized in Table 13.

Cross section calculations

Theoretical predictions of the single top cross section have been performed at next-to-leading order [73] for *s*-channel and *t*-channel diagrams. (Associated W production is not expected to be observable at the Tevatron, due to the massive final state (a W boson and a top quark) that must originate from a bottom sea quark.) These calculations were performed with two regularization techniques—phase space slicing and massive dipole subtraction—to remove infrared divergences and give consistent results. Loop uncertainties were calculated with three different renormalization methods and results were shown to be inconsistent only in finite terms connected to infrared divergences. These calculations are used to extract the theoretical cross sections for single top production.

The numbers used for the cross sections were derived by Sullivan [74] using the program ZTOP [75], assuming proton-antiproton collisions with center-of-mass energy 1.96 TeV. Sources of theoretical uncertainty come from uncertainty on PDFs, the mass of the bottom quark, the value of the strong coupling constant, the choice of renormalization scale, and the mass of the top quark, which for this calculation was taken to be 175 ± 4.3 GeV/c^2.

The *s*-channel calculations are similar to Drell-Yan production, which is well-studied. Leading-order kinematic distributions match next-to-leading order distributions very well when scaled by a *k*-factor of 1.54. The final result, including both top and antitop production, is an inclusive cross section of 0.88 ± 0.12 pb. (Top and antitop production are the same at the Tevatron.)

The *t*-channel diagram takes the analytic form of double deep inelastic scattering (DIS). The light quark probes a proton with DIS scale of Q^2, which is the virtuality of the W boson. The fermion line containing the top quark has a DIS scale of $Q^2 + m_t$. Using these scales results in kinematic distributions that do not change between leading order and next-to-leading order. The final inclusive cross section for *t*-channel is 1.98 ± 0.28 pb.

7.2 QCD multijet estimate

Having a model of the QCD multijet sample allows a method of estimation of its rate by fitting to a kinematic distribution. Missing transverse energy is a natural choice because QCD multijet events dominate the region with little missing transverse energy, since they have no true neutrino. Removing the missing transverse energy cut creates a sample with a large QCD multijet component which can be fit to the data. The samples described above are fixed to their expected value while the normalization of the QCD multijet and W + jets samples is fit to the data. The fraction of QCD multijet events that pass the missing

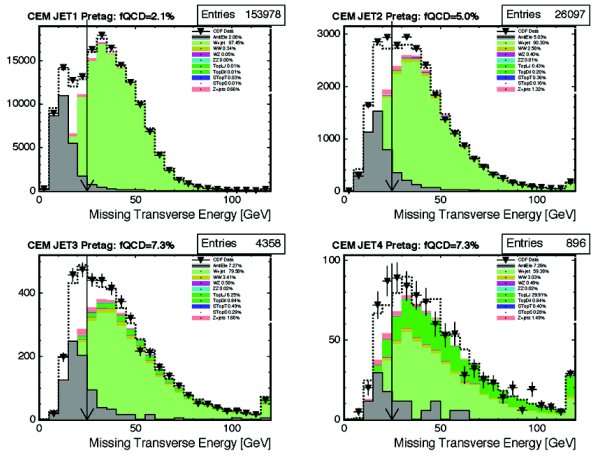

Figure 41. Fits to the missing transverse energy distribution to obtain the QCD multijet fraction in the pretag sample. The plots show events with one through four jets for the central electron trigger.

transverse energy cut can than be calculated. This is done before the tagging requirement is applied in order to get a proper estimate of the W + jets normalization, described next.

After the W + jets normalization is calculated, there remains the question of determining the QCD multijet fraction in the final sample after the b-tag requirement is applied. This is done by performing the same fit again in the tagged sample, giving the final estimate of the QCD multijet fraction in the sample.

Because of differences in the detectors and backgrounds, each lepton type is fitted separately. The untriggered muons require 35 GeV of missing transverse energy to the trigger and thus might seem to have no region of low missing transverse energy. However, the trigger calculation of missing transverse energy does not include the contribution of the

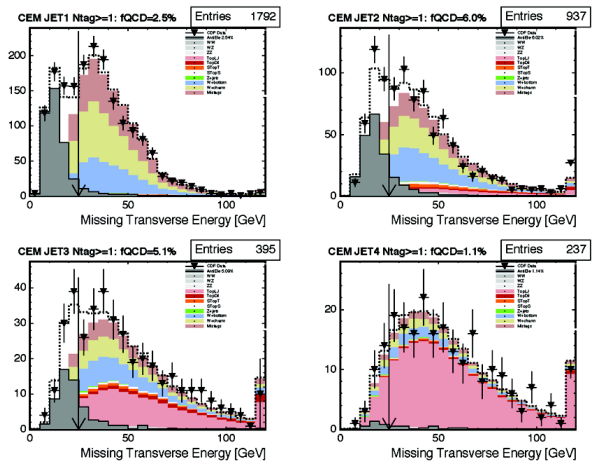

Figure 42. Fits to the missing transverse energy distribution to obtain the QCD multijet fraction in the tagged sample. The plots show events with one through four jets for the central electron trigger.

muon energy. Once this correction is performed, there is a substantial contribution of events with low missing energy, making the fit possible.

This fraction can fluctuate substantially in different regions or with different choice of histogram binnings. In addition, the sample itself may be mismodeled and its low statistics affect sensitivity because the estimate relies on the high-missing-transverse-energy regime of the distribution, which can have very few events in it. A large systematic uncertainty of 40% covers all observed effects and leaves some room for some more. In the case of double-tagged events, there is so little data and so few expected QCD multijet events that the fits are very uncertain, and an 80% uncertainty is applied.

7.3 W + heavy flavor

The W + jets cross section cannot be easily calculated for several reasons. One problem is the large number of possible diagrams; another is that the diagrams contain many instances of gluons that split to produce two quarks. This causes an infrared divergence when the angle between them is small. This can be controlled by applying a factorization scale which defines the energy at which two jets are considered as a single jet. The value of this factorization scale, which is a mathematical artifact and has no intrinsic physical meaning, has a substantial impact on a leading-order cross section calculation, and thus the absolute cross section calculated by ALPGEN is not trustworthy. Furthermore, ALPGEN only calculates cross sections to leading order, which underestimates the next-to-leading-order cross section. Another method is necessary to extract this contribution.

If the contribution of all other backgrounds is accounted for, any remaining events must be from the W + jets sample. Using this assumption, the W + jets contribution can be estimated directly from the data. It is safer to use the *pretag* sample—events which pass all selection cuts except the b-tag requirement, which is not yet applied—to ensure that this estimate is statistically independent of the signal sample. This sample is ten to twenty times larger than the final sample, and it is dominated by light jets. This gives an estimate for the overall normalization of the W + jets sample.

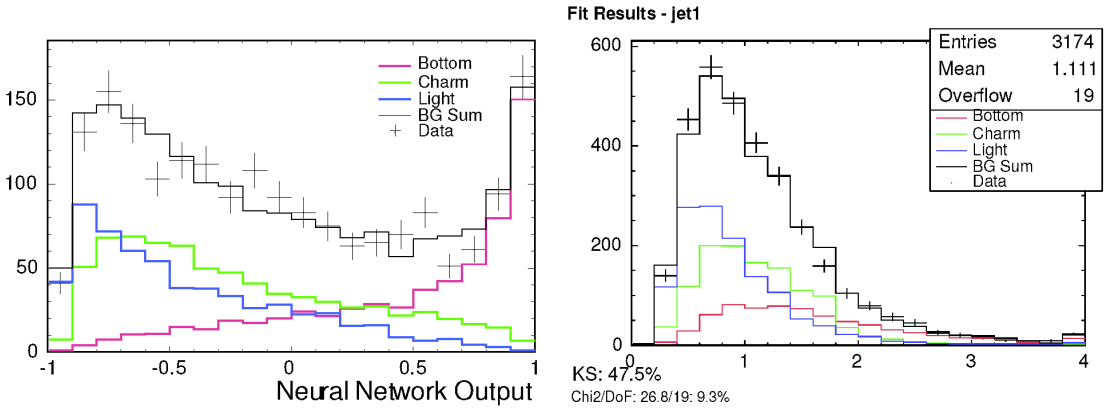

Figure 43. The results of fits to flavor-sensitive distributions to calibrate the $W+$ heavy flavor fraction in the one-jet tagged events: (left) the output of a neural network which separates jet flavors, and (right) the invariant mass of the secondary vertex, using the loose definition to increase its statistical power.

It is necessary now to distinguish the components of the sample including bottom and charm quarks, which form the largest part of the background. This heavy flavor fraction is calculated by looking directly at Monte Carlo events, divided and weighted as described previously. Then the tagging rate for each sample is calculated using the same method as for the Monte Carlo samples. This gives an estimate of the rate of $W+$ heavy flavor jets in the tagged sample.

However, the Monte Carlo simulation does not properly predict the heavy flavor fraction in this sample. It requires an additional factor to match the data. This factor is calculated in the statistically independent sample which has only one jet. This sample has high statistics and is dominated by $W+$ jets events and a negligible contribution of single top

events. This makes it ideal to estimate the heavy flavor content. To do this, the Monte-Carlo-based and QCD multijet background contributions are subtracted from the data, and the remainder is fit to the W + jets templates in a way similar to the determination of the mistag parameter α. Two flavor-sensitive variables are used: the secondary vertex mass of the loose b-tagger, and a neural-net jet-flavor separator.

The loose b-tagger has an advantage in that it has more events because it has a higher acceptance than the tight tagger used by default. However, its discrimination between heavy and light flavors is not as good as the neural-net flavor separator, which will be described in detail on page 158. Both quantities are measured, and the k-factor is assigned as a weighted average of the fit values for b and c quarks in both distributions. The systematic error, taken to be large enough to include all data points as well as unity, giving 1.4 ± 0.4. A cross-check in the two-jet bin shows that the fit values are well covered by this uncertainty.

The problem with the heavy flavor estimate is assumed to have to do with the difficult theoretical problem of the infrared divergence in the case of gluon splitting. Thus, the contribution of the $W + c$ + jets sample, which has no gluon splitting, is not scaled; only events with two heavy quarks in the final state are scaled by this factor. A recent measurement of the $W + c$ + jets cross section at CDF [76] is consistent with the ALPGEN calculation, reinforcing this view.

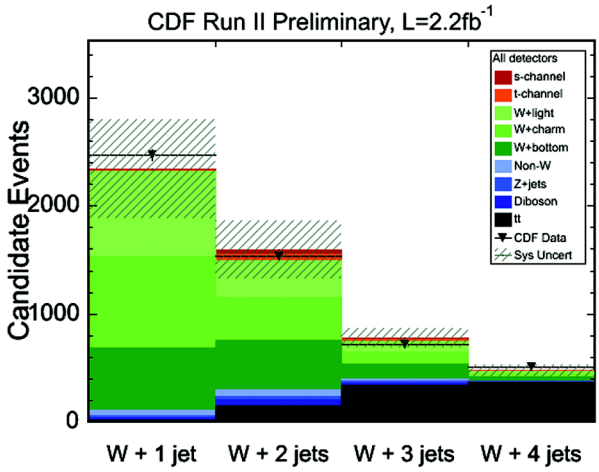

Figure 44. The predicted number of events as a function of number of jets. This shows a good agreement between the prediction and the data; it also shows that the single-top signal is far too small to be seen by a simple counting experiment.

Because the systematic uncertainty on the heavy flavor fraction is so large, and because the fit adjusts the fraction to match the data, no additional systematic uncertainties (such as amount of radiation, factorization scale, or choice of PDF) should be required for this number. However, the uncertainty is rounded up to be 30% of the final sample in order to cover any small contributions which may have been missed.

7.4 Mistags

Once the heavy flavor has been properly estimated, the remainder of the pretag sample is assumed to be W + light jet events. This is by far the largest contribution to the pretag sample, which means that a substantial number of events will be mistakenly tagged.

The mistag rate is estimated by applying the mistag probability to the W + jets sample: multiplying each event by its mistagging probability allows an estimate of the number that make it into the final sample.

The uncertainty on this estimate comes from the mistag matrix, which has uncertainty in each bin because of the finite statistics of the sample it was derived from. Most of the uncertainty comes from the uncertainty on the parameter α.

7.5 Event yield

Table 14 shows the final background estimate in the different signal regions of the analysis. The uncertainties on the estimates include the systematic uncertainty on the mistag matrix, the k-factor, the QCD multijet estimate, the b-tagging scale factor, the lepton ID efficiency scale factor, the primary vertex position scale factor, and the trigger efficiency. They also include the uncertainty on the luminosity obtained and, in the case of Monte-Carlo-derived estimates, and the cross section calculation.

Process	Two-jet events		Three-jet events	
	Single *b*-tag	Double *b*-tag	Single *b*-tag	Double *b*-tag
W + bottom	407.2 ± 122.7	54.4 ± 16.9	121.1 ± 36.5	20.0 ± 6.2
W + charm	398.7 ± 120.2	5.3 ± 1.7	106.1 ± 32.6	2.8 ± 0.9
W + light quark	338.4 ± 55.7	1.4 ± 0.4	100.8 ± 16.7	1.0 ± 0.3
Diboson	59.6 ± 5.9	3.7 ± 0.5	20.0 ± 2.0	1.4 ± 0.2
Z + jets	25.4 ± 3.8	1.2 ± 0.2	10.2 ± 1.5	0.7 ± 0.1
Top pair	119.2 ± 16.8	26.9 ± 4.4	261.5 ± 36.4	77.7 ± 12.8
QCD multijet	58.4 ± 23.4	1.5 ± 0.6	22.5 ± 9.0	0.2 ± 0.1
Total background predicted	1398.0 ± 250.9	94.2 ± 19.4	642.2 ± 79.9	103.9 ± 14.5
s-channel	31.8 ± 4.4	9.4 ± 1.5	10.2 ± 1.4	3.3 ± 0.5
t-channel	60.4 ± 8.8	1.7 ± 0.3	15.8 ± 2.3	2.5 ± 0.4
Total signal predicted	92.2 ± 13.2	11.1 ± 1.8	26.0 ± 3.7	5.8 ± 0.9
Events observed	1434	101	595	117

Table 14. The expected event yield for the four different samples used in this analysis, given in number of events. The background prediction is consistent with the number of data events; however, the systematic uncertainty on the background prediction is far larger than the expected single-top signal.

Chapter 8

The matrix element method

The background estimate (Table 14) shows that the expected single top signal is substantially smaller than the systematic error on the background estimate. This means that it is impossible to obtain a significant result by simply counting events. A more sophisticated approach is needed.

The previous generation of published single-top searches [77] performed a likelihood fit to sensitive distributions, such as H_T, the scalar sum of all jets, the lepton, and the missing transverse energy in the event, which is sensitive to the mass peak in single-top-quark diagrams; or $Q*\eta$, the lepton charge multiplied by the pseudorapidity of the untagged jet, which exploits the kinematic feature of *t*-channel processes that the light quark tends to have large pseudorapidity and its direction is correlated with the lepton charge [78]. However, while this approach improves the sensitivity markedly over simple event counting, it is still sub-optimal: there are several sensitive kinematic distributions, so fitting only one of them excludes information from another one.

The matrix element method takes a different approach: rather than searching for sensitive variables, it starts at the fundamental prediction from quantum field theory and uses

the differential cross section calculation to quantify how likely an event is to come from a given process [79].

8.1 The differential cross section

The basis of the matrix element calculation is Fermi's golden rule, derived earlier and repeated here for convenience: for scattering of two particles with four-momenta q_1 and q_2 into n particles with four-momenta p_i, the differential cross section is given by

$$d\sigma = |M|^2 \frac{\hbar^2 S}{4\sqrt{(q_1 \cdot q_2)^2 - (m_1 m_2 c^2)^2}} d\Phi_n$$

where S is a combinatoric factor for identical particles, m_i is the particle mass,

$$d\Phi_n = \delta\left(q_1 + q_2 - \sum_{i=1}^{n} p_i\right) \prod_{i=1}^{n} \frac{c d^3 p_i}{(2\pi)^3 2E_i}$$

is the phase space factor, and M is the matrix element for the interaction [80].

This can be used to calculate the likelihood that a given set of four-vectors of final-state particles resulted from a given interaction. If all information about initial- and final-state particles were known—including particle type, spin, color, energy, and momentum—and if the matrix element for the interaction could be calculated exactly, then this differential cross section would give all possible information about how likely a given

event is to come from a given process. Comparing differential cross sections of different processes for each event would result in a perfect analysis, extracting all possible information from each event and comparing it to theory.

Unfortunately, practical details make the actual calculation more difficult. Because the matrix element cannot be exactly calculated, because some particles cannot be measured, because partons cannot be measured directly but only after showering and hadronization, and because of the finite resolution of the components of the CDF detector, many assumptions must be made in order to complete the calculation, each assumption causing a commensurate loss in sensitivity.

8.2 Parton distribution functions

If all final-state momenta and energies are fully known, the initial-state momenta and energies can also be calculated from conservation of energy and momentum, under the assumption that the initial partons have no transverse momentum. (In reality, they have a small transverse momentum, but this is a relatively small effect.) However, the likelihood of a given initial-state configuration depends on the parton distribution functions (PDFs), which give the probability of a given momentum for a given type of parton. This scale varies by the momentum transfer Q^2, which is different for each process. The calculation of this value is not clearly defined because Q^2 is an artifact of perturbative calculation and cannot be measured directly.

The Q^2 calculation is chosen to be the same as that used for the theoretical calculation. For *s*-channel, following Sullivan [81], the process works similarly to Drell-Yan scattering, which uses $Q^2 = \hat{s} = \left|q_1 + q_2\right|^2$ for initial quark four-vectors q_1 and q_2. The top pair diagram, having the same event topology as *s*-channel, uses the same calculation. For *t*-channel, a double deep inelastic scattering calculation gives the momentum difference for the light quark line, and the momentum difference plus the square of the top quark mass for the line including the top quark. All W + jets diagrams use $m_W^2 + \left(\sum_{jets} p_T\right)^2$.

The addition of this calculation changes the differential cross section to

$$d\sigma = |M|^2 \frac{\hbar^2 S}{4\sqrt{\left(q_1 \cdot q_2\right)^2 - \left(m_1 m_2 c^2\right)^2}} f\left(q_1\right) f\left(q_2\right) d\Phi_n$$

where the *f*s are the PDFs of the incoming proton and antiproton. For diagrams in which the initial-state quarks are different—the *t*-channel and *Wc* diagrams—it is not clear which initial-state particle came from the proton and which from the antiproton, so both combinations are evaluated and their probabilities are added.

8.3 Phase space

The phase space factor includes differential terms for the momentum of each particle divided by the energy of the particle and a constant factor. This gives

$$d\sigma = |M|^2 \frac{\hbar^2 S}{4\sqrt{(q_1 \cdot q_2)^2 - (m_1 m_2 c^2)^2}} f(q_1) f(q_2) \frac{c d^3 p_l}{(2\pi)^3 2E_l} \frac{c d^3 p_\nu}{(2\pi)^3 2E_{\nu i}} \prod_{i=1}^{n_{jets}} \frac{c d^3 p_i}{(2\pi)^3 2E_i} \delta^{(4)}\left(q_1 + q_2 - p_l - p_\nu - \sum_{i=1}^{n_{jets}} p_i\right)$$

This can be simplified by considering the input quarks to be massless. Also, constant factors can be removed; later it will be shown that any constant factors will drop out of the final result (see page 167). This gives

$$d\sigma = |M|^2 \frac{f(q_1) f(q_2)}{|q_1||q_2|} \frac{d^3 p_l}{E_l} \frac{d^3 p_\nu}{E_\nu} \prod_{i=1}^{n_{jets}} \frac{d^3 p_i}{E_i} \delta^{(4)}\left(q_1 + q_2 - p_l - p_\nu - \sum_{i=1}^{n_{jets}} p_i\right)$$

8.4 Transfer functions

Unfortunately, none of the quantities in the equation can be measured directly, because the detector cannot measure individual partons. While leptons can be measured well, jets resulting from quarks and gluons are much harder to measure, and neutrinos cannot be detected at all by the CDF detector. This problem is addressed by *transfer functions*, functions which map between the energies and momenta of final state particles and objects actually observed in the detector. For transfer functions W the result is

$$d\sigma = \int |M|^2 \frac{f(q_1) f(q_2)}{|q_1||q_2|} W_l^3\left(p_l, p_{l\,meas}\right) \frac{d^3 p_l}{E_l} W_\nu^3\left(p_\nu, p_{\nu\,meas}\right) \frac{d^3 p_\nu}{E_\nu} \prod_{i=1}^{n_{jets}} W_i^3\left(p_i, p_{i\,meas}\right) \frac{d^3 p_i}{E_i} \delta^{(4)}\left(q_1 + q_2 - p_l - p_\nu - \sum_{i=1}^{n_{jets}} p_i\right)$$

where W^3 refers to the three transfer functions that are needed to map the energy, polar angle, angle, and azimuthal angle of the parton to their measured quantities.

The measurement of lepton quantities is assumed to be good enough that a Dirac delta function can be used. The same is true of the jet angles. While these assumptions are not perfectly true, they only reduce the potential sensitivity of the analysis; they do not affect the accuracy of the result. However, the jet energies are known to be very difficult, and they require transfer functions.

The jet energy transfer functions are modeled exactly as the jet energy correction for calorimeter response: a double Gaussian in the difference of the energies, linearized in the jet transverse momentum, requiring a total of ten parameters. Parton-level quantities are matched to reconstructed jets in Monte Carlo events and the transfer function is derived from a likelihood fit to extract the ten parameters. In principle, these transfer functions are the same as the jet energy corrections, except that jet energy corrections perform a correction based on the most probable value of the parton, while transfer functions are integrated to extract all information about the energies.

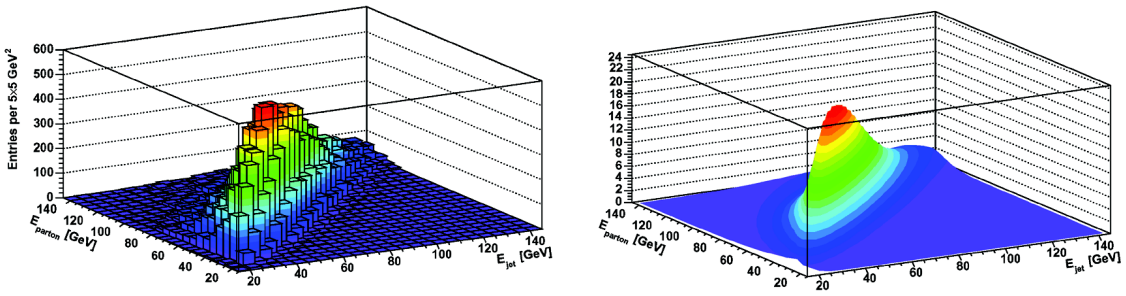

Figure 45. (top) The distribution of parton energy versus jet energy in Monte Carlo events, and (bottom) the parameterization with a double Gaussian distribution as a transfer function.

However, transfer functions will be different in different regions of the detector. Because of this, three sets of transfer functions are derived for three different regions of pseudorapidity: $|\eta| < 0.8$, $0.8 < |\eta| < 1.2$, and $|\eta| > 1.2$.

In addition, jets calculated with a cone of 0.4 miss some particles that would otherwise be included. Larger cone sizes will result in better energy resolution. Generally, increasing the cone size reduces signal acceptance by too often combining two partons into a single jet. However, in roughly 80% of two-jet events, all jets reconstructed with a cone of 0.4 can be matched to jets reconstructed with a cone of 0.7. In such events, the cone-0.7 jets give better energy resolution. Thus, two sets of transfer functions are derived, one for cone-0.4 jets and one for cone-0.7 jets. The cone-0.7 transfer functions are only used for events in which both jets can be matched to cone-0.7 jets.

Since *b*-quark jets, light-quark jets, and gluon jets all have different kinematic features, a separate transfer function is derived for each. The *b*-quark jet functions are

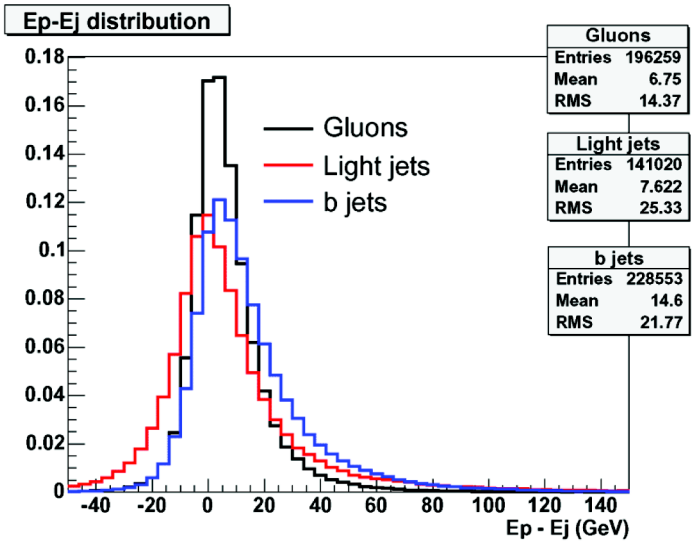

Figure 46. The difference in parton-level and jet-level energy for different kinds of jets, motivating the need for different transfer functions for each type.

derived from the bottom quark from the top quark decay in *s*-channel events; the light-quark jet functions are derived from the light quark in *t*-channel events; and the gluon jet functions are derived from the radiated gluon in Wc + jet events.

The x and y components of the neutrino momentum, and the z component of the momenta of the initial quarks (assuming no initial transverse momentum), all are derived from the conservation of energy and momentum. The longitudinal component of the momentum of the neutrino cannot be measured in the detector because the initial-state partons' longitudinal momentum is not known; each parton is part of a proton or antiproton with longitudinal momentum 980 GeV/c, but the fraction carried by each quark is not known, so the total longitudinal momentum is not known. This can be dealt with by

integrating over the neutrino's longitudinal momentum, allowing it to be as large as the beam energy (within the laws of conservation of energy and momentum).

This changes the differential cross section to

$$d\sigma = \int dp_{z_\nu} |M|^2 \frac{f(q_1) f(q_2)}{|q_1||q_2|} \prod_{i=1}^{n_{jets}} \frac{dE_i W(E_i, E_i)}{E_i} \frac{\delta^{(4)}\left(q_1 + q_2 - p_l - p_\nu - \sum_{i=1}^{n_{jets}} p_i\right)}{E_l E_\nu}$$

where all other parton-level quantities have been replaced by their measured values. For a given set of measured quantities, the probability density can be calculated by numerical integration.

8.5 Matrix element

The matrix element of any scattering process cannot be calculated exactly because no analytic form for its solution exists. For the practical purposes of this analysis, the matrix element can only be calculated to leading order, because the loop corrections involved in higher-order calculations are computationally prohibitive and often do not noticeably affect the kinematic properties of the event.

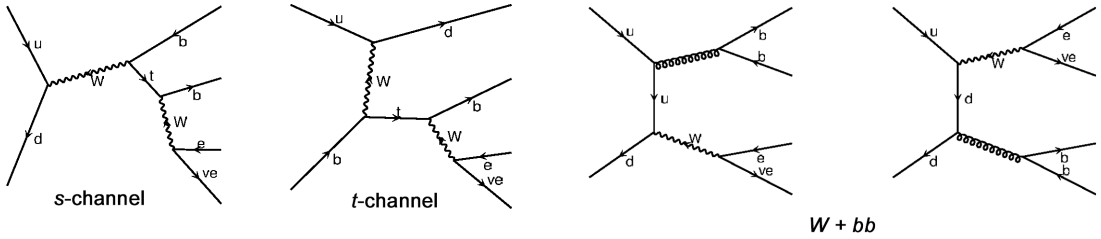

Figure 47. Feynman diagrams used to calculate *s*-channel, *t*-channel, and *W* + *bb* probabilities for two-jet events.

The matrix elements are obtained from MadGraph [82], the same program used to generate signal Monte Carlo events. MadGraph uses a library called HELAS [83] to calculate leading-order matrix elements; this code (rewritten in C++ for performance reasons) is used to calculate the matrix element for each event.

Matrix elements are calculated for each signal and major background process: *s*-channel single top, *t*-channel single top, *Wbb*, *Wc*, *Wgg*, and top pair production. The *Wc* and *Wgg* diagrams are not computed in the three-jet case because they do not contribute significantly to the background of three-jet events. The other backgrounds either cannot be associated with a single matrix element (as in the case of QCD multijet events) or are only small contributions, so that adding their matrix elements does not increase sensitivity (as in the case of diboson processes). However, it is not necessary to include all background processes; the purpose of the calculation is to separate signal events from background events, and while leaving out a matrix element may affect sensitivity, it does not affect the accuracy of the answer.

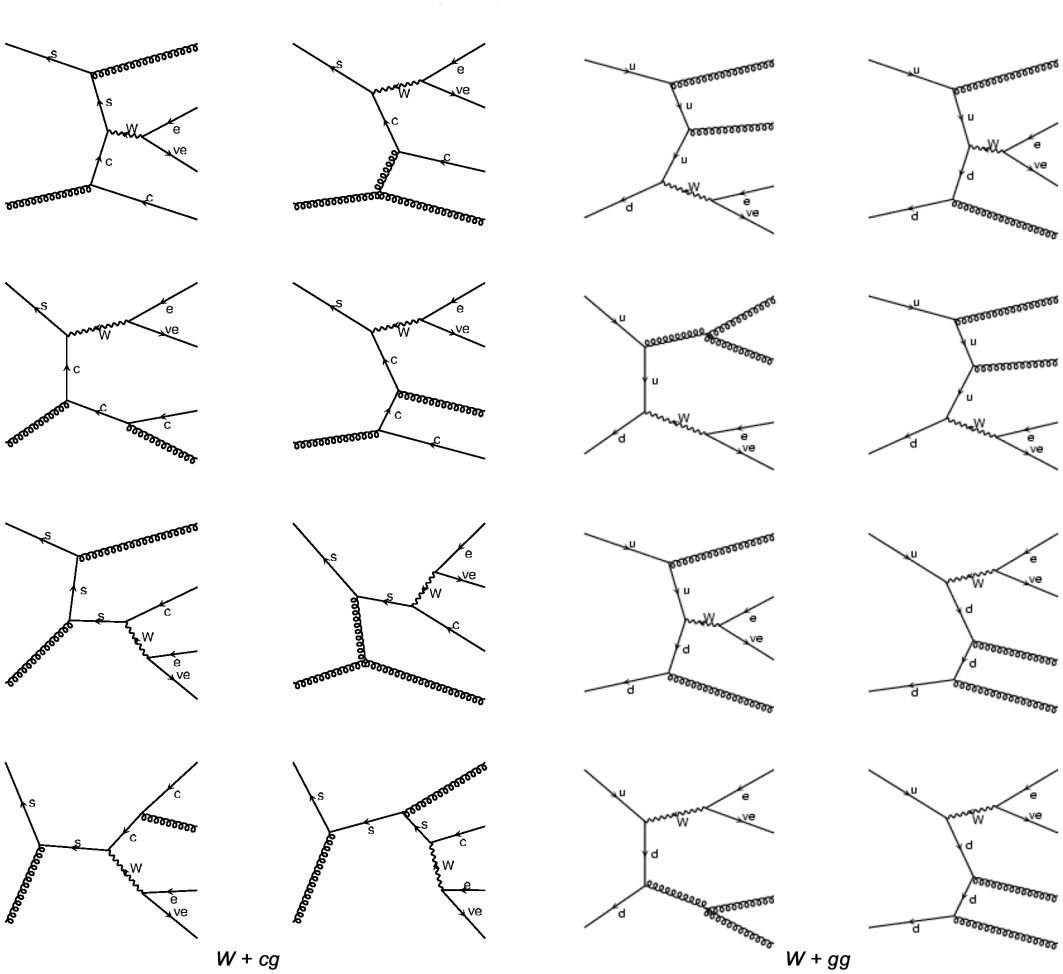

Figure 48. Feynman diagrams used to calculate $W + c$ and $W + gg$ probabilities for two-jet events.

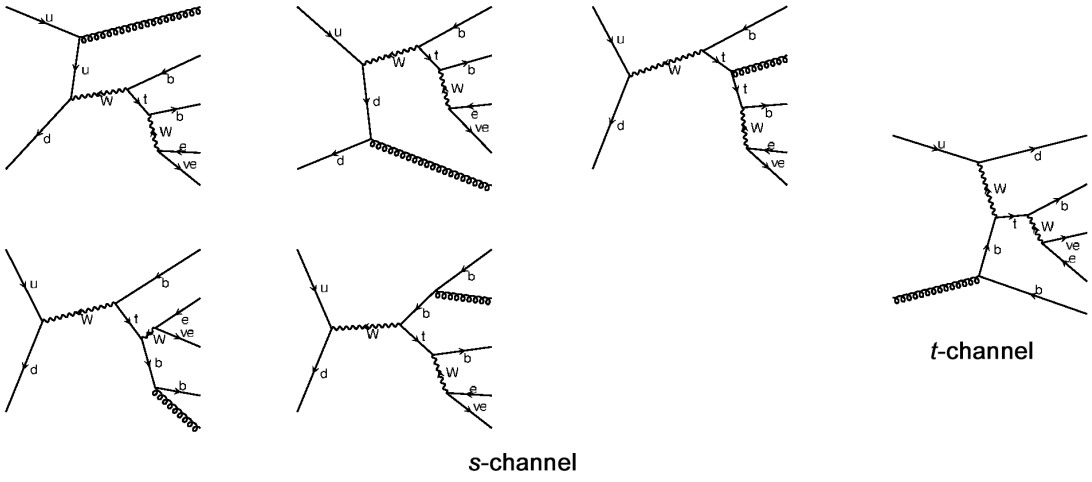

Figure 50. Feynman diagrams used to calculate *s*-channel and *t*-channel probabilities for three-jet events.

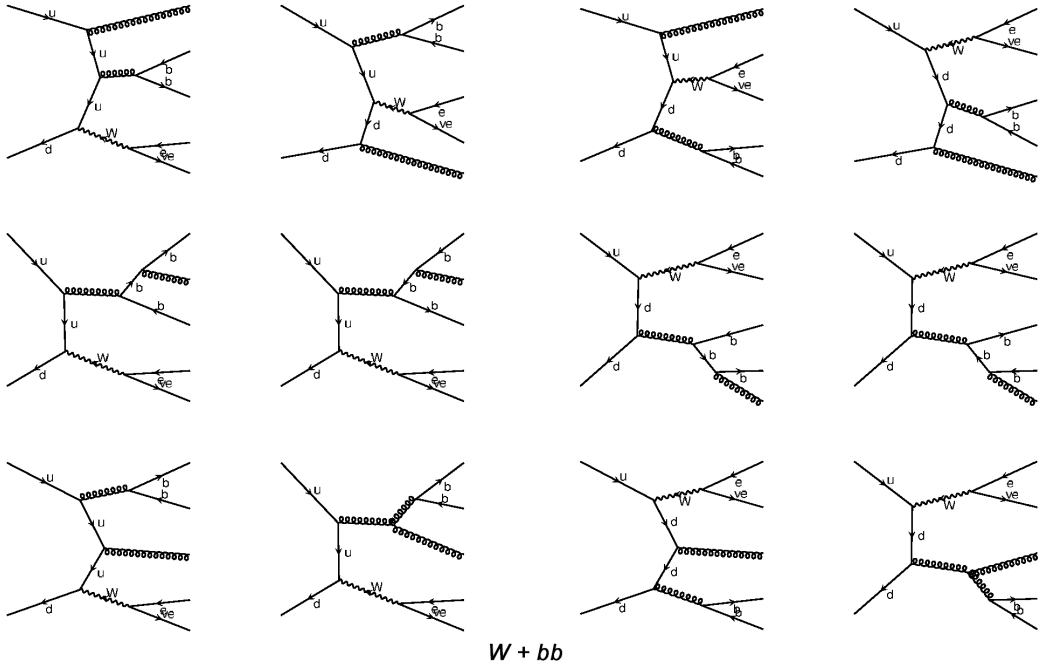

Figure 49. Feynman diagrams used to calculate $W + bb$ probabilities for three-jet events.

Top pair production matrix element

The diagram for top pair production poses a special problem for the matrix element method because its final state is not the same as that of single-top events. The only way a top pair event can be accepted as a two-jet event is if some final state particles are missed: either both W bosons decay leptonically and one lepton goes undetected, or one W boson decays hadronically and two of the final-state partons are undetected. Because both of these scenarios can occur a significant fraction of the time, this case requires special care.

In the two-jet case, the problem is solved by using a diagram in which one W boson decays leptonically and the other is treated as a final-state particle that is not observed in the detector—whether it decays leptonically or hadronically, it is not observed. All three components of the momentum of this final-state W boson are then integrated over all possible momentum configurations. This allows the calculation of the top pair production matrix element, though three extra integrations must be performed.

In the three-jet case, the diagram used assumes that one W boson decays to leptons and the other decays hadronically, and that one of the light quarks is missed. Its momentum is integrated in the same way, also adding three integrations.

8.6 Combinatorial issues

Several of the diagrams have ambiguities in their final state: for example, the s-channel has two kinematically distinct bottom quarks in the final state. Choosing which jet

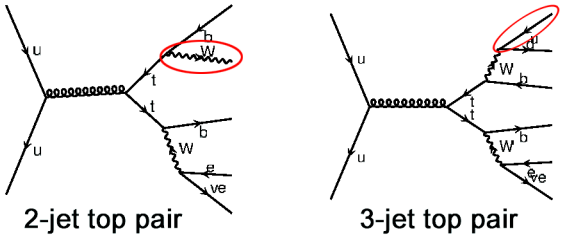

Figure 51. Feynman diagrams used to calculate top pair probabilities for two- and three-jet events. The circled particles are assumed to be unobserved and an integral is taken over their momenta.

should be matched to a given parton is difficult; this analysis solves the problem by calculating the differential cross section under both assumptions and adding the answers together. However, in the case of the *t*-channel diagram, the tagging information is used to match the final-state bottom quark to the tagged jet, improving the sensitivity of the calculation.

The combinations of matching jets to quarks are chosen based on the principle that heavy quarks should be matched to tagged jets whenever possible. In cases of ambiguity, all different combinations are tried. This is true except in the case of the top pair diagram, which has too many ambiguities and is too computationally expensive to try all combinations. In this case, only the two combinations of assigning tagged jets to bottom quarks are calculated.

8.7 Numerical integration

The differential cross section calculated from the matrix element method must be calculated for every event, including both data and Monte Carlo events. This requires an integral over the neutrino's longitudinal momentum, the energies of all jets, and, in the case of top pair production, the momenta of missing particles. This results in a three- to seven-dimensional integral. It is not possible to perform this integral analytically, so a numerical method is used.

For the simpler integrals in the case of *s*- and *t*-channel signal and *Wbb* background in two-jet events, integration is performed using the adaptive quadrature [84] method based on the CERNLIB [85] RADMUL [86] routine, adapted for Root [87], then adapted again for this analysis. The basic idea of this algorithm is to divide the *n*-dimensional integration region into equal-sized regions and estimate the uncertainty in each one. The region with the largest uncertainty is divided in half. This continues until all regions have a smaller error than that requested (1% in this analysis), at which point the value of the integral in each region is estimated at the sum is returned.

The advantages of adaptive quadrature are its stability and its reproducibility: unlike some integration routines, it does not rely on random numbers and thus is fully deterministic. An *s*-channel matrix element calculation converges to the desired accuracy in an average of about five seconds in this analysis' implementation. However, for larger integrals, such as

the six-dimensional calculation required by the top pair production diagram in two-jet events, this algorithm becomes prohibitively slow—in a test, the calculation was unable to converge on a single integration even after an entire day of computation. A faster integration routine is clearly needed.

The algorithm chosen is DIVONNE, based on CERNLIB's DIVON4 [88] function, as implemented in the CUBA integration library [89]. DIVONNE is a Monte-Carlo-based integrator using stratified sampling to subdivide its regions. Stratified sampling minimizes the variance of the Monte Carlo events thrown in a given subregion. The Koksma-Hlawka inequality [90] shows that the variance is bounded by half the volume of the subregion times the difference of the supremum and infimum of the function in that subregion. The borders of the subregion are adjusted to reduce this spread. Once a requested variance is reached, the integral is estimated by adding the total of randomly generated points in each subregion.

The implementation of the algorithm in CUBA also samples the subregions independently with the same number of Monte Carlo events in each region. If this result is not consistent with the integral derived already, the regions are subdivided further and the process is repeated.

The DIVONNE algorithm gave results consistent with RADMUL when tested on an ensemble of a thousand events. It is also very stable: running it repeatedly on an identical event was never observed to change the result by more than 0.001%. The top pair integral that RADMUL was unable to perform was evaluated by DIVONNE in about five minutes.

In this analysis, most differential cross section calculations take between one and ten seconds, except for the top pair matrix element, which takes an average of five minutes to calculate both combinations. The use of the DIVONNE algorithm (plus careful optimization of the integration loop) makes it possible to calculate differential cross sections for the millions of Monte Carlo events in the various samples available. The total computing time used is still not small—close to a million CPU-hours—but feasible given a few months of computation on good processing farms.

Chapter 9

Analysis

Each event, in both data and Monte Carlo, is processed by the matrix element calculation. The differential cross sections for each signal and background process are calculated and stored for each event. These numbers, which in principle include all kinematic information about the event, are used to separate the single top production signal from the background.

9.1 Jet flavor separator

Unfortunately, some detector-specific information is not carried into the matrix element calculation. In particular, *b*-tagging is far from perfect, and many jets that are tagged do not come from bottom quarks. Being able to distinguish mis-tagged jets makes it much easier to distinguish the signal, which is characterized by bottom quarks, from some of the W + jets background.

This analysis uses a jet flavor separator [91] based on an *artificial neural network*. Artificial neural networks attempt to emulate the functionality of the human brain to solve a variety of problems, particularly those involving pattern recognition. Rather than requiring users to calculate and program the intricacies of a particular pattern, an artificial neural

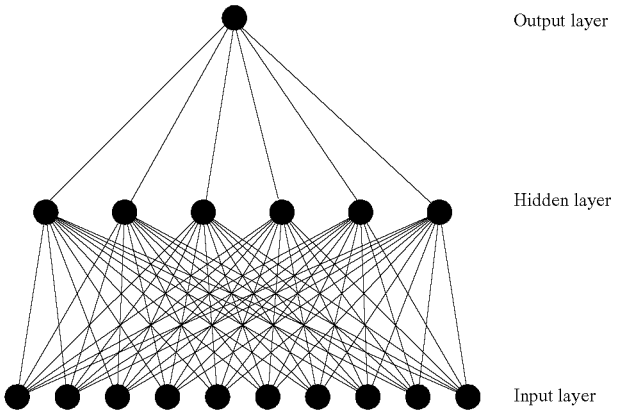

Figure 52. A diagram of a neural network. Input variables are put in nodes at the input layer. Each line represents a weight applied by the network, and each node produces a single output described by a sigmoid function. This network produces a single number between –1 and 1 at the output node. The weights are chosen during the training of the neural network.

network is *trained* by being given different events to examine, letting the network calculate the necessary interrelations to distinguish the different cases. In this analysis, several variables are given as inputs to the network, which returns a single value between –1 and 1, where –1 is background-like and 1 is signal-like. In this case, *b*-jets are considered to be signal and other jets are considered as background.

The artifical neural network used for the jet flavor separator was constructed using the NeuroBayes [92] program. This includes a process of Bayesian regularization [93] to prevent "overtraining"—the possibility of finding a correlation from a statistical fluctuation in the training sample that does not exist in the data. It also calculates the relative

importance of a given variable in separating signal from background and removes variables that do not significantly improve separation.

The artificial neural network is constructed in three layers: an input layer, a hidden layer, and an output layer. The input layer contains all input variables as well as a bias node, which is used to adjust the relative contribution of each input variable. Each of these nodes is connected to each node in the hidden layer, and each node in the hidden layer is connected to a single output node (Figure 52).

Each node returns the weighted sum of its inputs, where the weights are quantities adjusted by the training of the network. This sum is transformed by a *sigmoid* function, $S(x)=\frac{2}{1+e^{-ax}}-1$, which forces the output value to be between –1 and 1. The network adjusts the weights to minimize the *entropy error function* $E_D=\sum_i \log\left(\frac{1}{2}\left(1+T_i\right)o_i+\varepsilon\right)$, where T is –1 for background and 1 for signal, o is the output of the neural network function, and ε is a regularization parameter which solves some numerical subtleties and is gradually reduced to zero. This function characterizes the difference between the network's output and the true value T of each event. Minimizing this function is performed by the method of gradient descent [94], in which each weight is adjusted by the amount suggested by its approximated first derivative at a given point.

This training is repeated many times, and each time the training program intentionally introduces some statistical fluctuations. This causes statistically significant structures to intensify with more iterations, while statistical fluctuations are reduced through

Input variables used in neutral-network jet flavor separator

- Number of tracks with impact parameter significance greater than 3
- Signed impact parameter significance of the three highest-momentum tracks
- Whether the first pass of the SECVTX algorithm was successful
- Invariant mass of the secondary vertex
- Transverse momentum of muon in a jet with respect to the jet axis
- Number of good tracks in the jet
- Pseudo-$c\tau$ of secondary vertex
- Number of tracks used in the vertex fit
- Transverse energy of the jet
- Pseudorapidity of second highest-momentum track with respect to the jet axis
- Momentum of the three highest-momentum tracks
- Significance of three-dimensional vertex displacement
- Transverse momentum of the highest-momentum track with respect to the jet axis
- Transverse energy of electron in a jet
- L_{xy} significance
- χ^2 per degree of freedom of vertex fit in the first pass
- Total transverse momentum of tracks in secondary vertex in the first pass
- Total transverse momentum of tracks in secondary vertex divided by transverse energy of jet
- Impact parameter significance of third-highest momentum track
- Impact parameter of second- and third-highest momentum tracks
- Number of muons in the jet
- Transverse momentum of a muon in the jet
- Total charge of tracks in second-pass vertex fit

Table 15. The variables used in the jet flavor separator to distinguish *b*-jets from other jets.

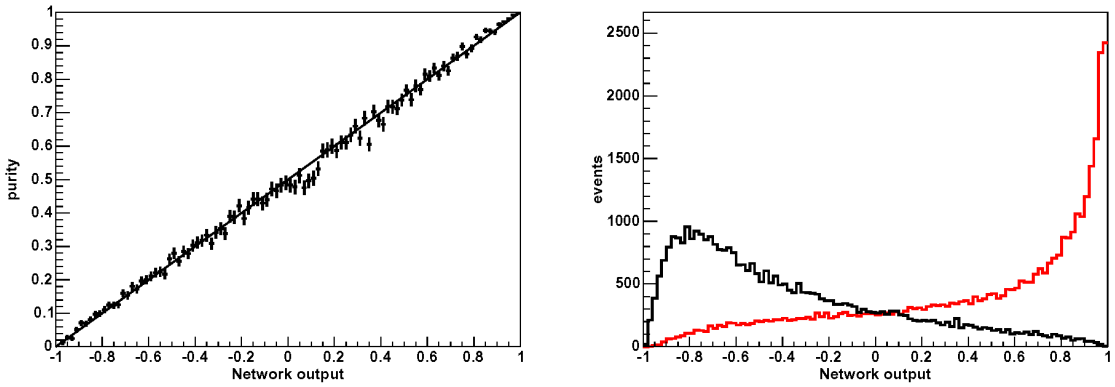

Figure 53. (left) Purity of the sample as a function of neural network output. (right) *b*-jet signal (red) and light-jet background (black) shapes of the output of the neural network after training.

an effect known as *weight decay*. Weights which have no statistical significance are gradually reduced, and they are removed when they become insignificant. This process is repeated until the network stabilizes.

The artificial neural network used for the jet flavor separator was trained on jets from Monte Carlo W + jets events, requiring each jet to have a transverse energy greater than 20 GeV and to be tagged by the secondary vertex tagger. Jets that were matched to underlying bottom quarks are taken as signal and other jets are treated as background. The network uses 31 variables, which are listed in Table 15. These variables are properties of the secondary vertex and its tracks that allow a good separation of jets of different flavors (Figure 53). The output of this network has a linear slope in *b*-jet purity, which means it can also be treated as the probability that a single jet comes from a bottom quark.

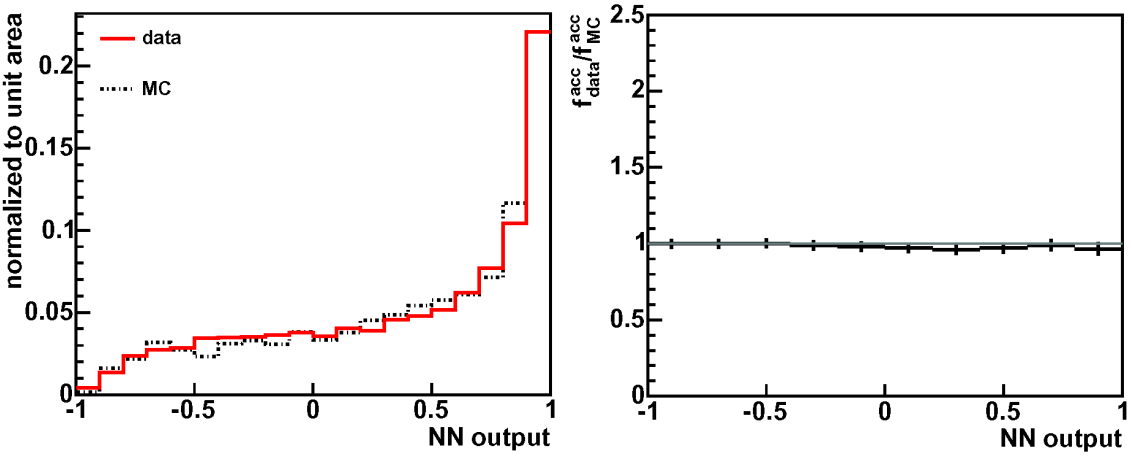

Figure 54. Comparison of data and Monte Carlo in *b*-jet-enriched sample. (left) Neural network output in data and Monte Carlo, and (right) ratio of cumulative acceptance as a function of neural network output, showing that data and Monte Carlo are consistent.

Once the network is trained, it is compared to data, looking at both *b*-jet-enriched and mistag-enriched samples. The *b*-jet-enriched sample is the same as that used by the electron method of the *b*-tagging scale factor calculation (see page 110): two jets, one of which is tagged and one of which has a high-momentum electron. The distribution of the neural network output is found to match very well with the Monte Carlo templates (Figure 54).

The mistag sample is examined by looking at events with negative tags, as with other mistag estimates (see page 113). Because there is a significant disparity between the data and the Monte Carlo in this sample (Figure 55), a correction function is applied to Monte Carlo events to make the output distribution match the data. This function has a dependence

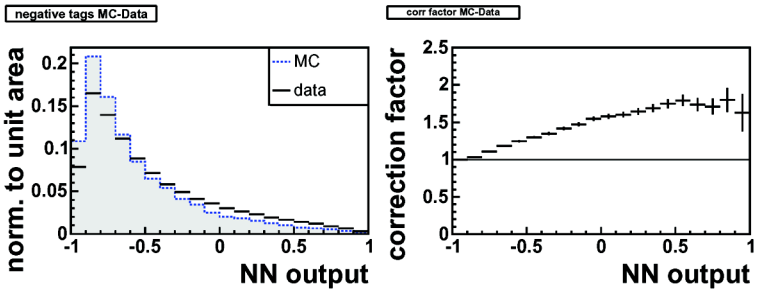

Figure 55. (left) Distribution of negative tagged events in data and Monte Carlo and (right) ratio of their cumulative acceptances. This shows the need for a correction function.

on jet transverse energy, the number of tracks in the vertex, and the sum of the transverse energies of the tracks in the vertex.

The artificial neural network was tested on all signal and background processes and shown to not have a significant dependence on sample type; that is, *b*-jets from top pair events look very similar in the neural network distribution to *b*-jets from $W + bb$ events. This shows that the network can be applied safely applied to any sample without correction; it looks only at process-independent jet information.

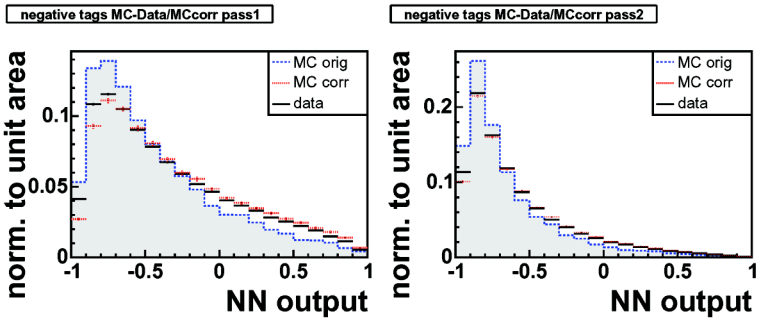

Figure 56. Neural network output for negative tagged jets before and after correction for (left) events tagged in the first pass of SECVTX, and (right) events tagged in the second pass.

This powerful neural network variable poses a problem when it is applied to the W + light jets Monte Carlo sample. In this sample, most events have no b-tagged jet (they are instead weighted by the mistag matrix), and so they have no secondary vertex and a jet-flavor separator value cannot be calculated. Because of this, these events are assigned a random value taken from the distribution of W + light jet events in Monte Carlo.

The case of QCD multijet events, which cannot be b-tagged because of insufficient statistics, is even more complicated because a substantial number of real bottom quarks is expected in this sample. This is resolved with the neural-network jet flavor separator: templates of the three flavors are constructed and fit to the data in the QCD-multijet-enriched control region with missing transverse energy between 15 and 25 GeV. This gives a proportion of approximately 45% bottom quarks, 40% charm quarks, and 15% light quarks. Each QCD multijet event is triplicated, each is assigned a probability at random from the distributions of bottom, charm, or light quarks, and each is weighted according to the expected probability.

9.2 Event probability discriminant

Once the differential cross sections and jet-flavor separator values are calculated for each event, a single variable is constructed that discriminates between signal-like and background-like events. This is called the *event probability discriminant* (EPD), the variable used to extract the cross section in this analysis.

The output of the matrix element calculation is, up to some normalization factor, a differential cross section. For historical reasons involving the original application of this method [95], these outputs are usually referred to as *probabilities*, although in this analysis they are not properly normalized. The question of whether these numbers should be called probabilities or likelihoods is largely semantic: these numbers are the output of a differential cross section calculation, and they have no statistical meaning beyond being the output of a function which is used to construct the EPD variable. For the rest of this paper, in deference to tradition, these numbers will be referred to as probabilities.

Simply, the EPD is the signal probability divided by the sum of the signal and background probabilities. This is motivated by the Neyman-Pearson lemma [96], which says that a likelihood ratio is the most sensitive variable for separating hypotheses. This gives a distribution that is bounded by zero and one, with a very background-like event being at zero and a very signal-like event being at one. The fact that the EPD is a ratio has the added effect that systematic effects common to both signal and background probabilities will cancel out to first order. To add the flavor separator information, each matrix element with a final-state *b*-jet is multiplied by the *b*-jet probability from the neural network (scaled to fall between zero and one), while each matrix element without a final-state *b*-jet is multiplied by one minus the *b*-jet probability:

$$EPD = \frac{b\left(P_{schan} + P_{tchan}\right)}{b\left(P_{schan} + P_{tchan} + P_{Wb\bar{b}} + P_{t\bar{t}}\right) + (1-b)\left(P_{Wc\bar{c}} + P_{Wc} + P_{Wgg}\right)}$$

The $W + cc$ probability is taken to be the same as the $W + bb$ probability (since the diagram differs only in the masses of the final-state quarks), but it is written separately because the tagging probability changes for that assumption (the $W + cc$ matrix element has no final-state b-jet). In the case of double-tagged events, the b-jet probability b is replaced by the product $b_1 b_2$ for matrix elements with bottom quarks in the final state, and $(1-b_1)(1-b_2)$ for the other matrix elements.

Each of these probabilities has a hidden normalization factor that can be considered part of the probability calculation. In theory, this number accounts for the relative contribution of this probability to the overall separation: the total cross section of the process combined with the effects of the relative efficiency of the event selection cuts;however, these effects are so convoluted and difficult to calculate that these factors are better derived pragmatically. These coefficients shift the distribution of the EPD, changing the percentage of signal events in each bin. In this analysis, the coefficients are adjusted to optimize sensitivity: they are chosen to minimize the average uncertainty on the single-top cross section in many pseudo-experiments (see page 181).

The coefficients are adjusted separately in each of the four channels in which this

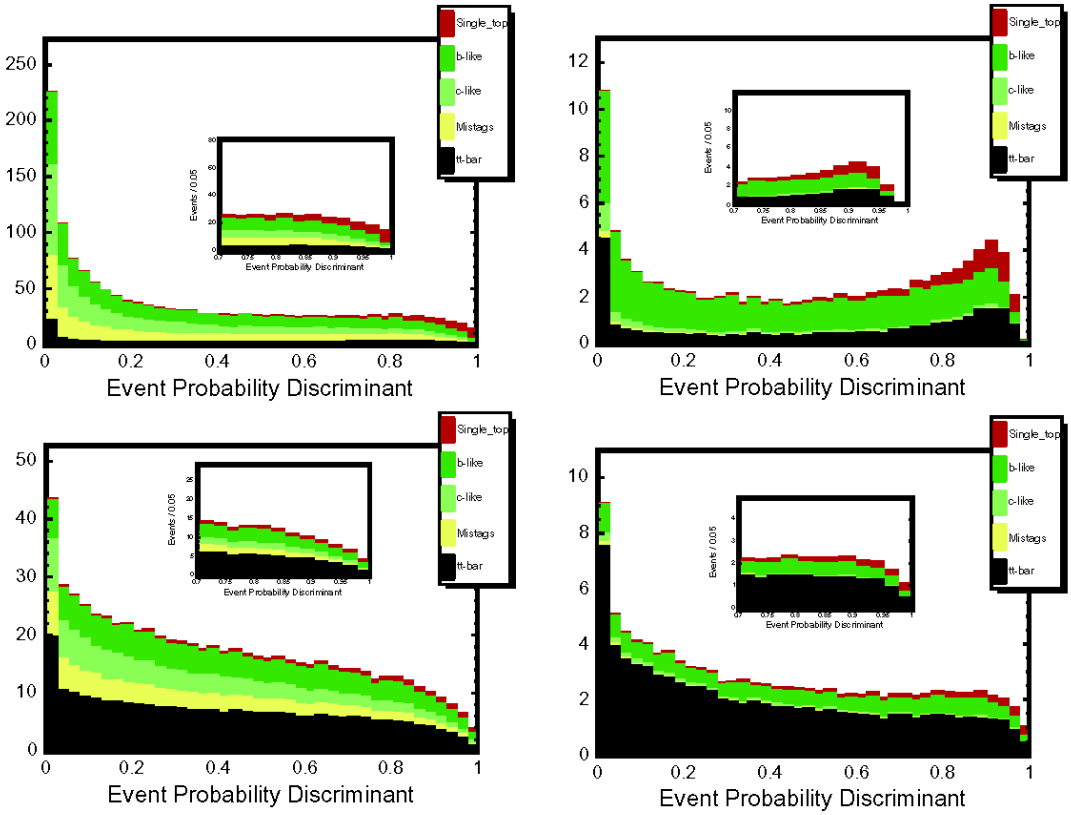

Figure 57. Templates of EPD for events with (top left) two jets and one *b*-tag, (top right) two jets and two *b*-tags, (bottom left), three jets and one *b*-tag, and (bottom right) three jets and two or more *b*-tags. For ease of display, the templates have been grouped together: *b*-like templates include the $W + b$ template as well as the WZ, ZZ, Z + jets, and QCD multijet templates; *c*-like templates include the $W + c$ and WW templates.

analysis is performed: events with two or three jets and with one or two tagged jets. (The extremely small triple-tagged sample is included with the double-tagged sample.) Then histograms of the EPD for each Monte Carlo sample are constructed for use as templates (Figure 57). The number of bins in the histogram is optimized to maximize the expected sensitivity of the analysis; however, the binning was found to have little effect, so near-

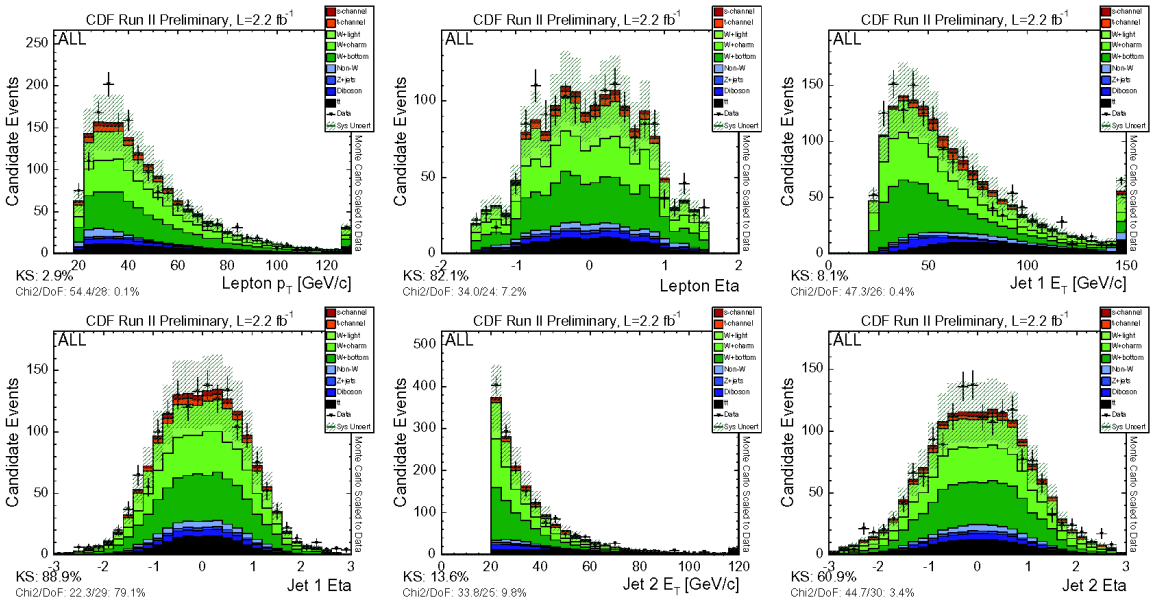

Figure 58. Some of the validation plots for the two-jet sample with at least one *b*-tag.

optimal sensitivity is reached with histograms of forty bins. At this point, measuring the cross section is simply a matter of comparing the data to the Monte Carlo prediction.

9.3 Cross-checks

Before performing the cross section measurement, it is important to be sure that the Monte Carlo events properly describe the data. This is done by constructing a set of *control plots* to examine different quantities and make sure that the data are statistically consistent with the Monte Carlo prediction. This is done in many different kinematic variables, including pseudorapidity and transverse momentum distributions of leptons, jets, and missing transverse energy, as well as the angles between them.

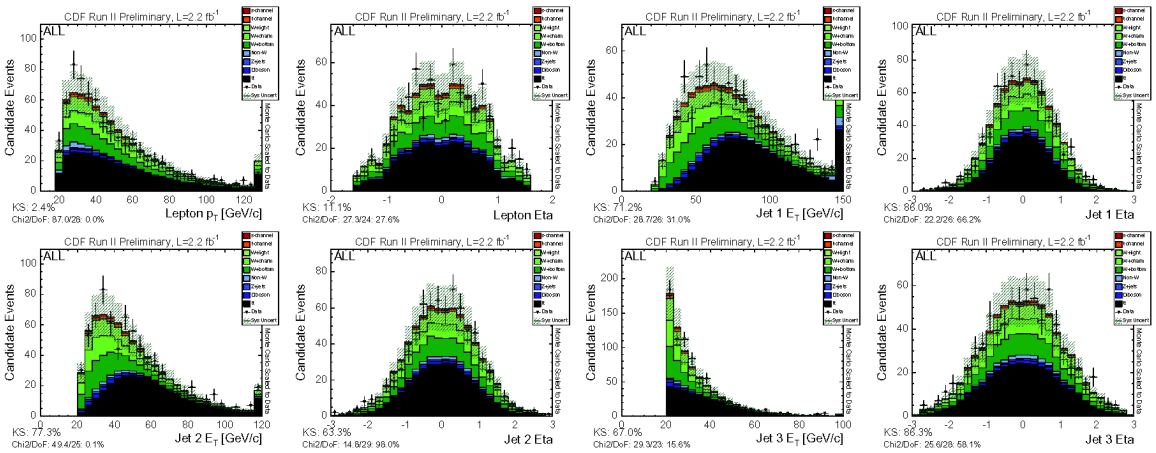

Figure 59. Some of the validation plots for the three-jet sample with at least one *b*-tag.

The agreement between the data and Monte Carlo events can be quantified by one of two statistics. One is the $\chi^2 = \sum_{bins} \frac{N_{expected} - N_{observed}}{\sigma}$ statistic. The χ^2 distribution can be used to calculate, for a given number of degrees of freedom, the probability that a given set of data comes from a given distribution. The drawbacks of this method are that it does not take systematic uncertainties into account and it does not include bin-to-bin correlations. It is possible to construct a modified χ^2 statistic that properly deals with these problems, but such calculations are computationally intensive and rarely add interesting information. This analysis approximates the uncertainty σ with the square root of the number of events observed.

The other test statistic is the *Kolmogorov-Smirnov* (KS) *probability*. This calculates the cumulative sum of *n* histogram bins, starting with the left-most bin, and returns the largest value of the cumulative difference between data and Monte Carlo distributions as *n*

increases from 1 to the total number of bins in the histogram. Then a thousand pseudo-experiments are created by choosing a random number for each bin according to a Poisson distribution centered at the bin's expected value in Monte Carlo, and the KS statistic is calculated for each pseudo-experiment. The percentage of the time that the result is lower than the measured test statistic is reported as the KS probability, which is the probability that a data distribution results from a given distribution in Monte Carlo. The advantage of this technique is that is takes into account bin-to-bin correlations, since several consecutive bins in which the data are consistently higher than the prediction will cause the KS statistic to increase. It does not include systematic uncertainties, however, so it is most useful in regions in which the uncertainty is dominated by statistical uncertainty, as in the tagged sample.

The control plots show good agreement insofar as the limited data allow. However, the number of tagged events is not large enough to distinguish small (but potentially significant) differences between the data and Monte Carlo distributions. Thus it is useful to check these distributions in a statistically independent control region with a large amount of data. For this purpose, the untagged sample, which has no secondary vertex tags, is useful. This sample is dominated by W + light flavor events and has many thousands of events in data, but very few single top events (less than half a percent).

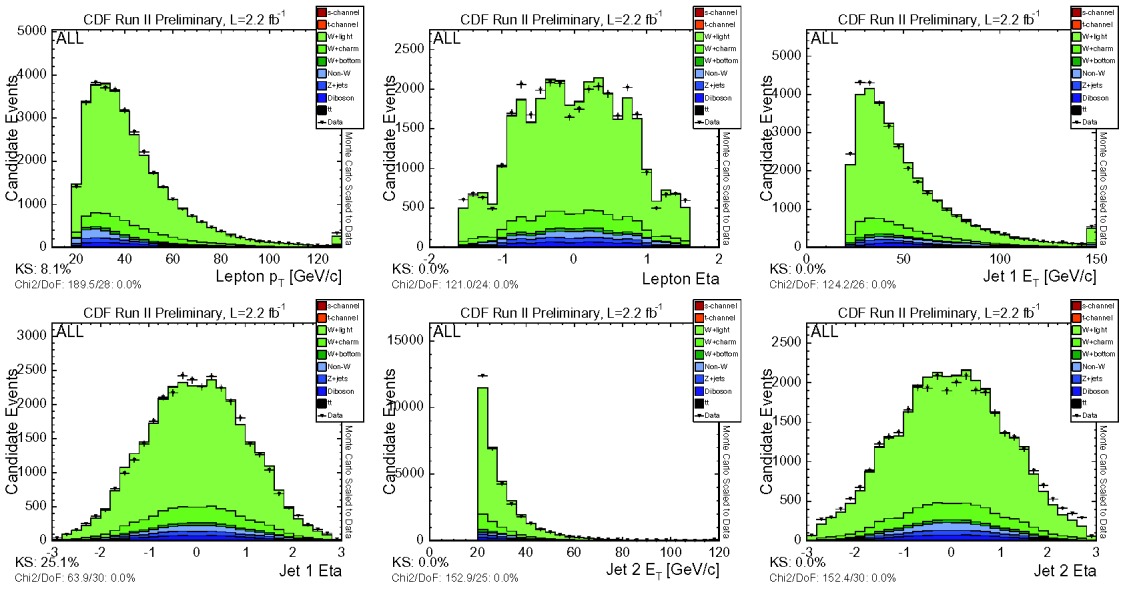

Figure 60. Some of the validation plots in the two-jet control sample with no *b*-tags.

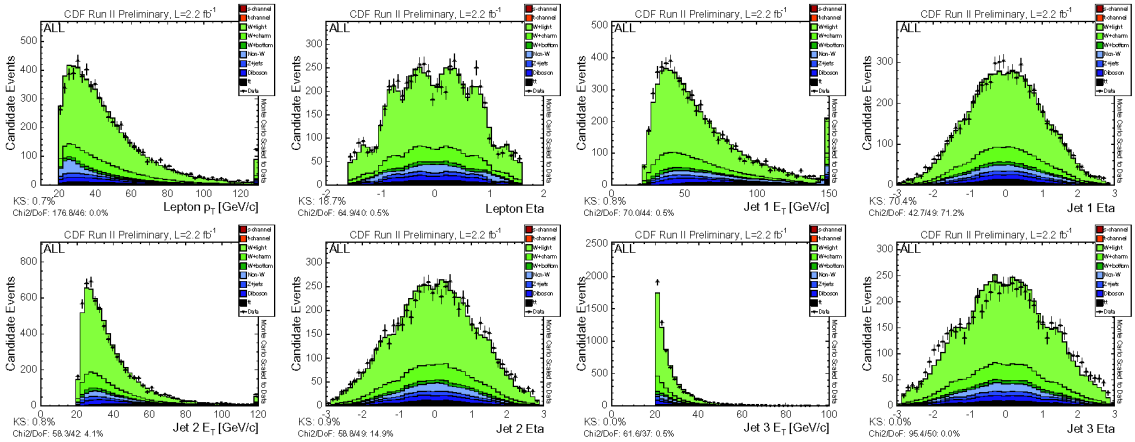

Figure 61. Some of the validation plots in the three-jet control sample with no *b*-tags.

The checks in the untagged region look generally good—except for two distributions, which will be addressed in the section on systematic uncertainties (see page 193). The χ^2

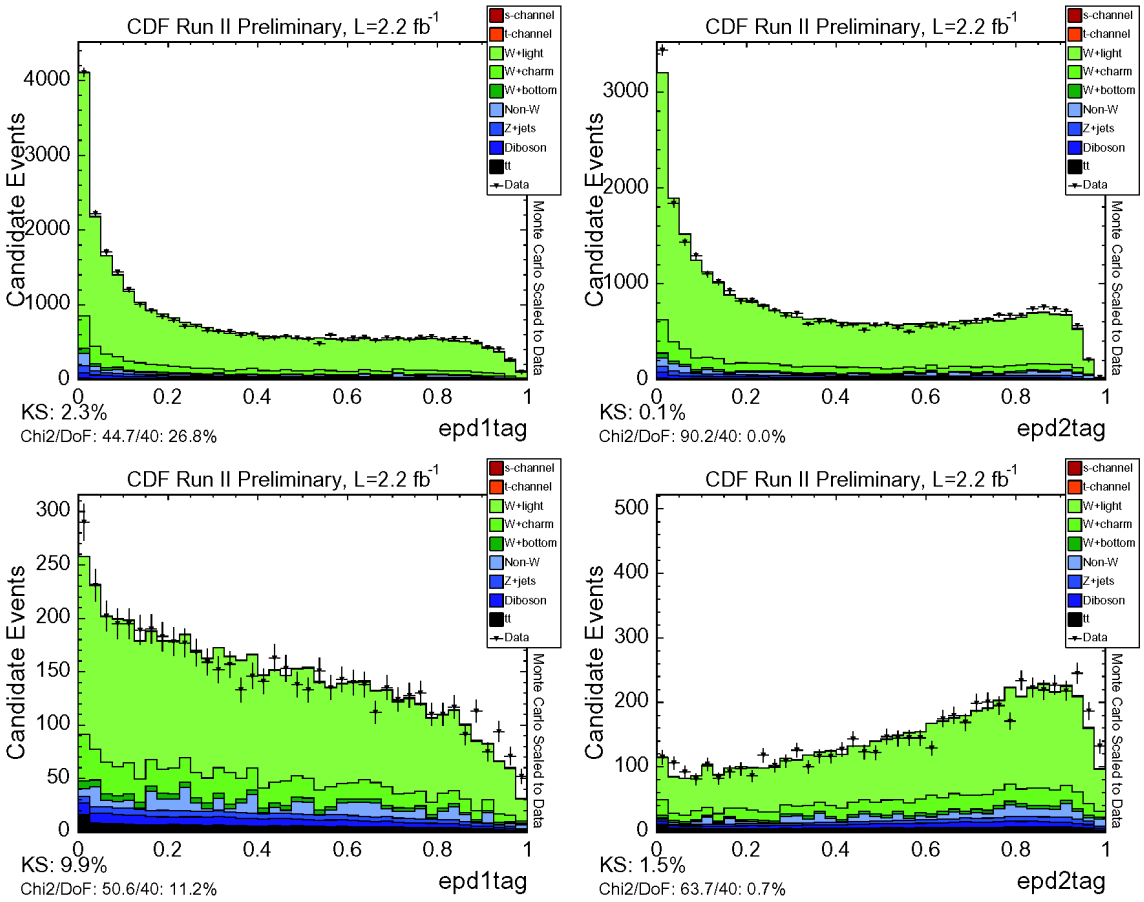

Figure 62. Distributions of EPD calculated in the untagged sample. The top row shows two-jet events and the bottom row shows three-jet events. Because the EPD calculation uses two different sets of coefficients for single- and double-tagged events, it is necessary to check both sets of distributions, even though there is no tagging information in this sample. The left-hand plots show the single-tag discriminant and the right-hand plots show the double-tag discriminant.

and KS probabilities are generally underestimated for these distributions because their uncertainties are dominated by systematic uncertainties rather than statistical uncertainties, and neither test statistic includes the effect of systematic uncertainties.

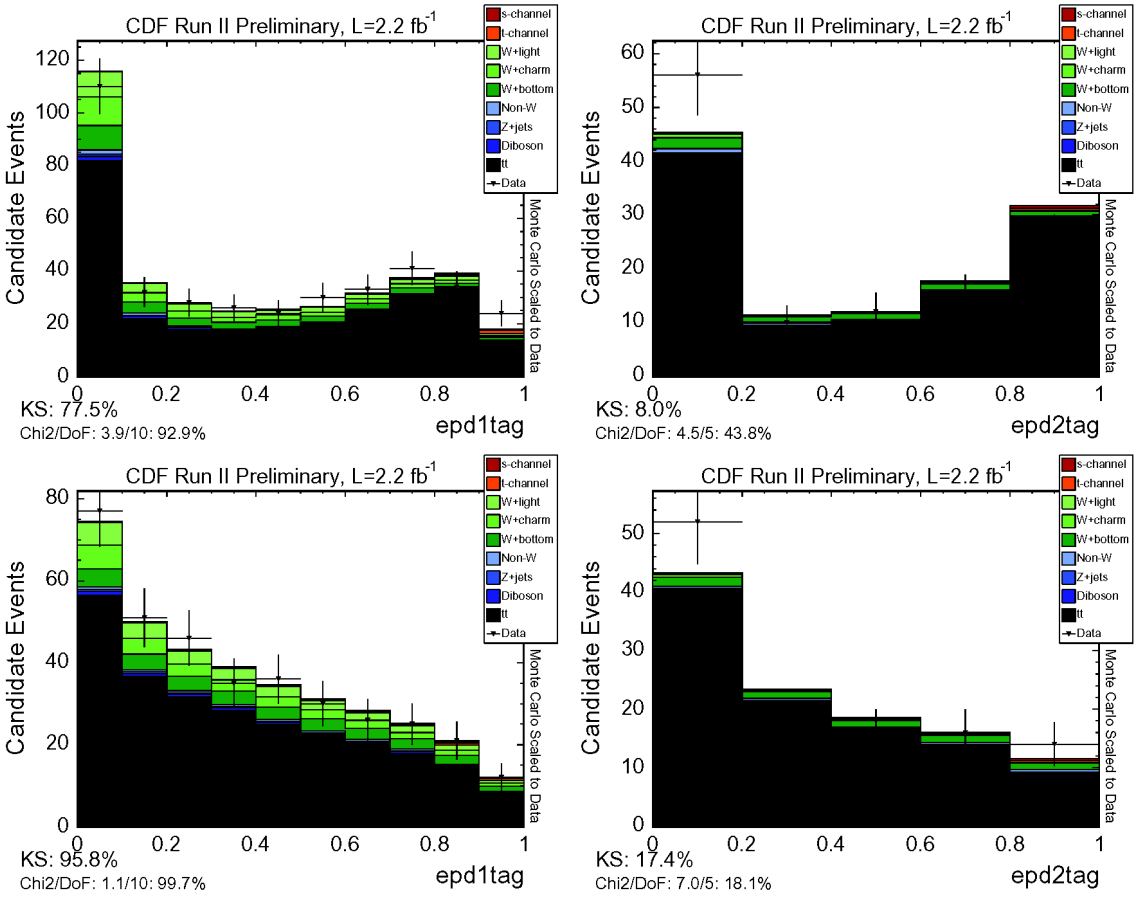

Figure 63. Distributions of EPD calculated in the four-jet sample. The top plots show the distributions for the two-jet assumption, with the two lowest-energy jets discarded; the bottom plots show the distributions for the one-jet assumption, with the lowest-energy jet discarded. The left-hand plots show events with one *b*-tag; the right-hand plots show events with more than one *b*-tag.

As an additional check, the EPD is calculated for untagged events, with all *b*-jet probabilities set to 0.5 (which causes them to cancel out of the equation). This is a check to make sure that Monte Carlo events behave in the same way as the data in this sample,

particularly testing the distribution of W + jets events, which dominate the sample. This cross-check shows an excellent match between data and Monte Carlo.

Another check is performed in the top-rich four-jet tagged sample. In these events, the jets with lowest transverse energy are removed to get the appropriate number of jets, and then the EPD is calculated. This checks the shape of the top pair production background in the tagged region, and complements the untagged check. This check also shows good agreement with the data.

9.4 Likelihood function

Now the question remains of how to extract the cross section for this measurement. This analysis uses a *Bayesian* [97] approach: calculate the *likelihood* that the data comes from the distributions predicted for different amounts of signal, and use this to construct a posterior probability density as a function of single top production cross section. The maximum value of this curve lies at the most probable value of the single top production cross section given the data, and the uncertainty is defined as the region around this value that contains 68% of the total area of the probability density curve. Bayesian measurements require a *prior* which represents the experimenter's belief in the distribution of values of the outcome before examining the result. This analysis uses a prior that is flat for all nonnegative values of the cross section.

All that remains is the construction of the likelihood function. In a simple counting experiment with one signal process, one background process, and no systematic uncertainties, this likelihood would simply be

$$L(\beta_s) = \frac{e^{-\mu}\mu^n}{n!}$$

where $\mu = \beta_s N_s + N_b$. N_s is the number of signal events expected from the Standard Model prediction, β_s is a fraction of the Standard Model expectation (β_s = 1 for the Standard Model value), N_b is the number of background events, and n is the number of observed data events. The x axis of the posterior probability density is β_s. This represents the inherent uncertainty of a measurement due to the statistical fluctuations of a quantum-mechanical process, which is described by a Poisson distribution.

Calculating the likelihood for a histogram instead of a simple counting experiment can be thought of as a combination of many statistically independent counting experiments, which makes the likelihood

$$L(\beta_s) = \prod_{k=1}^{N_{bins}} \frac{e^{-\mu_k}\mu_k^{n_k}}{n_k!}$$

where now $\mu_k = \beta_s N_{s\,k} + N_{b\,k}$ and k is the index of each bin. If there are several sources of background, then N_b is the sum of several background contributions. Furthermore, if there are several channels, then the bin index k can be defined to include the bins of all histograms.

The likelihood is complicated by the presence of systematic uncertainties. The simplest case involves uncertainties on the background estimate that are symmetrical. The Bayesian prescription for systematic uncertainties [98] is to treat them as *nuisance parameters*, imperfectly known numbers whose values affect the result but which are not themselves of interest to the analysis. Each nuisance parameter is assigned a prior that gives the degree of belief of different possible values and then *marginalized*, or integrated, to calculate the reduced likelihood as a function only of the parameter of interest (the cross section).

This analysis assigns all systematic uncertainties a Gaussian prior with a mean of zero and a width of one. In this case, zero represents the central value and ±1 represents the result of a shift up or down by the amount quoted as the uncertainty. This is a conservative treatment because most uncertainties are chosen to include at least 95% of the possible range of values, whereas treating an uncertainty as one standard deviation on a Gaussian covers only 68%.

Introduction of symmetric systematic uncertainties to a process requires integrations in the likelihood. For a single systematic uncertainty, parameterized by δ, which causes a percent change of ε in the background estimate, the likelihood becomes

$$L(\beta_s) = \int_{-\infty}^{\infty} d\delta \prod_{k=1}^{N_{bins}} \frac{e^{-\mu_k} \mu_k^{n_k}}{n_k!} \cdot \frac{e^{-\frac{\delta^2}{2}}}{\sqrt{2\pi}}$$

where now $\mu_k = \beta_s N_{s\,k} S_0 + \sum_{j=1}^{N_{proc}} N_{j\,k} S_j$ and $S_j = 1 + \delta\varepsilon_j$ is the systematic factor. The extra term in the likelihood is the Gaussian prior for the systematic uncertainty; the systematic factor adjusts the background for each template *j*, where *j* = 0 refers to the expected signal and other values of *j* refer to the different background processes.

If there are many systematic uncertainties, the resulting likelihood is their product:

$$L(\beta_s) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{N_{sys}} d\delta_i \prod_{k=1}^{N_{bins}} \frac{e^{-\mu_k} \mu_k^{n_k}}{n_k!} \cdot \frac{e^{-\frac{\delta_i^2}{2}}}{\sqrt{2\pi}}$$

where now $S_j = \prod_{i=1}^{N_{sys}} 1 + \delta_i \varepsilon_{ji}$. Here the systematic uncertainties are indexed by the variable *i*.

In some cases, the shifts caused by systematic effects are not symmetrical. In these cases, the prior used is two Gaussians connected by a Heaviside step function at zero. This gives a distribution that is discontinuous but whose effect on the output is continuous, since the effect of the systematic vanishes at zero. This prior gives equal probability to positive and negative values of the uncertainty. If each uncertainty on a given process *j* is given by a parameter $\varepsilon_{\pm}$, then the systematic factor changes to

$$S_j = \prod_{i=1}^{N_{sys}} 1 + |\delta_i| \left(\varepsilon_{ji+} H(\delta_i) + \varepsilon_{ji-} H(-\delta_i) \right)$$

This formula assumes that the $\varepsilon_{\pm}$ parameters have a sign included, for in this way it is possible to handle the case of uncertainties for which both positive and negative values of

the parameter result in an increase of the contribution of a given process.

Uncertainties on the background estimate, called *rate uncertainties*, are not the only possible uncertainties: there are also systematic uncertainties on the shapes of the Monte Carlo templates. These *shape uncertainties* can be calculated by comparing a Monte Carlo template with a shifted template and calculating the percent difference k between them in each bin; they can be considered as a bin-dependent rate uncertainty. Shape uncertainties are defined to change the shape of a distribution but not its overall normalization, which is controlled by the rate uncertainties. Thus each systematic effect may include either a rate uncertainty, a shape uncertainty, or both. This changes the systematic factor to:

$$S_{jk} = \prod_{i=1}^{N_{sys}} \left(1+\left|\delta_i\right|\left(\varepsilon_{ji+} H\left(\delta_i\right)+\varepsilon_{ji-} H\left(-\delta_i\right)\right)\right)\left(1+\left|\delta_i\right|\left(\kappa_{jik+} H\left(\delta_i\right)+\kappa_{jik-} H\left(-\delta_i\right)\right)\right)$$

Shape uncertainties cannot be reliably extrapolated beyond the point defined as one standard deviation, because that makes it possible for some bins to have a negative number of events. Thus, shape uncertainties are integrated between –1 and 1 only. In addition, some shape uncertainties are one-sided: that is, they define only a shift in one direction from the central value. For these uncertainties, the positive shifts ε_{ji+} and κ_{jik+} are given by the systematic, and the negative shifts are set to zero.

There remains one uncertainty not yet treated: the uncertainty in the prediction coming from finite Monte Carlo statistics. Because all templates are constructed from

Monte Carlo events, their predictions are never perfectly known, and for a rapidly falling distribution this may become a significant effect. To avoid overestimating the sensitivity of the analysis, a systematic uncertainty must be added which allows the prediction to fluctuate according to the number of Monte Carlo events in the bin.

The uncertainty on a prediction for Monte Carlo events is the square root of the number of events in a given histogram bin. However, since the events in question have weights attached to them (from ALPGEN scaling, mistag probability, trigger turn-on curves, and QCD multijet sample composition effects), the uncertainty is given by the sum in quadrature of the weights of the events in a given bin. Technically, these are separate nuisance parameters in each bin of each template; they can be dealt with by adding an integral to the equation. Thus, each bin content N_{jk} can be replaced by an additional integral with a Gaussian with a mean of N_{jk} and a width of σ_{MC}.

The final likelihood is :

$$L(\beta_s) = \int \cdots \int \prod_{i=1}^{N_{sys}} d\delta_i \prod_{k=1}^{N_{bins}} \frac{e^{-\mu_k} \mu_k^{n_k}}{n_k!} \cdot \frac{e^{-\frac{\delta_i^2}{2}}}{\sqrt{2\pi}}$$

where

$$\mu_k = \beta_s \tilde{N}_{sk} S_{0k} + \sum_{j=1}^{N_{proc}} \tilde{N}_{jk} S_{jk}$$

$$\tilde{N}_{jk} = \int_{-\infty}^{\infty} dl \frac{\left(N_{jk} + l\sigma_{MC}\right)}{\sqrt{2\pi}} e^{-\frac{l^2}{2}}$$

$$S_{jk} = \prod_{i=1}^{N_{sys}} \left(1 + |\delta_i| \left(\varepsilon_{ji+} H(\delta_i) + \varepsilon_{ji-} H(-\delta_i)\right)\right)\left(1 + |\delta_i| \left(\kappa_{jik+} H(\delta_i) + \kappa_{jik-} H(-\delta_i)\right)\right)$$

Coverage check

When using a Bayesian method, it is customary to check for *coverage* [99]: to show that 68% of all possible experiments fall within the quoted uncertainty. This is the *frequentist* definition of uncertainty, and showing coverage demonstrates that this analysis is insensitive to the choice of epistomological philosophy.

Performing a frequentist coverage check requires the generation of *pseudo-experiments*, sets of simulated data which indicate possible distributions of the true data. The basic method of creating pseudo-experiments is to assign, in each histogram bin of each template, a random number from a Poisson distribution with a mean centered at the total expectation of the Monte Carlo template. Adding the numbers obtained for each template gives an integral number of events in each bin that represents one possible fluctuation of the data set collected. To include systematic uncertainties, each of the systematic parameters is given a random number chosen from a zero-centered, unit-width Gaussian distribution. (Technically, this is a Bayesian modification to a frequentist method, but the frequentist

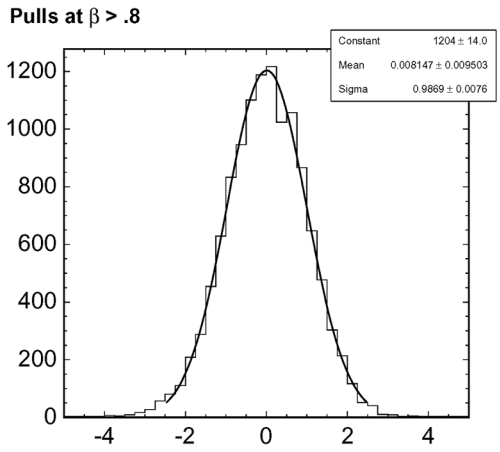

Figure 64. The distribution of pulls for many pseudo-experiments generated with signal fractions ranging from 0.8 to 4.0. A fit to a Gaussian distribution is applied to extract the mean and standard deviation.

prescription gives no guidance for the handling of systematic uncertainties.) In addition, each bin's prediction is altered by a random number chosen from a Gaussian distribution based on the number of Monte Carlo events in that bin, as described on page 180. After these systematic effects change the rates and shapes of the templates, random numbers from a Poisson distribution are chosen as before. These pseudo-experiments form an ensemble of distributions which simulate many different possible sets of data.

A convenient variable to check coverage is the *pull* of the measurement, defined as $\frac{\beta_{measured} - \beta_{expected}}{\sigma_{measured}}$. The pull has the advantage of being independent of the actual value measured. The distribution of pulls of many pseudo-experiments in an unbiased method with frequentist coverage should have a mean of zero and a standard deviation of one [100]. Figure 64 shows the distribution of pulls obtained for pseudo-experiments with single-top

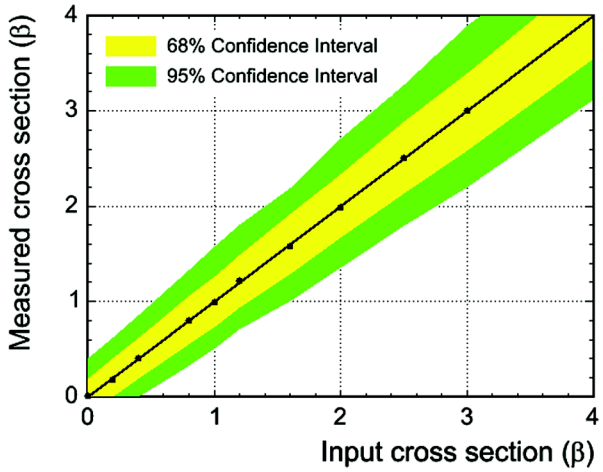

Figure 65. Linearity of the single top cross section measurement. Each point represents the median value of a thousand pseudo-experiments created with a given cross section. The black line represents a perfectly linear measurement.

contributions ranging from 0.8 to 4.0 times the Standard Model expectation. The mean and standard deviation show that this method has frequentist coverage and is unbiased.

Linearity check

Another important check is to make sure that in many pseudo-experiments, the median value measured is the same as the true value used to generate the pseudo-experiments. (The mean will be biased for small single-top contributions because the cross section is not allowed to be negative). To this end, sets of pseudo-experiments are created with single top values ranging from zero to four times the Standard Model expectation. Figure 65 shows the distribution of results for each point. The central point is the median value of the measured cross sections. The yellow and green bands show the region in which

68% and 95%, respectively, of all pseudo-experiments fall. This demonstrates that the method is linear and robust for different possible single-top contributions.

Sensitivity and significance

One common question for a measurement of this sort regards its significance: what does the measurement say about the existence of electroweak single top production? Is it possible that this resulted from a fluctuation from background, and that single top production does not occur? This is a question of *significance* which can be answered by computing the *p-value*. Pseudo-experiments are created assuming no single top production, and the *p*-value is the percentage of pseudo-experiments that give a result at least as signal-like as the measured data.

The choice of the test statistic to determine what "signal-like" means is arbitrary; however, the Neyman-Pearson lemma [101] indicates that the most sensitive test statistic is a likelihood ratio. Using the likelihood shown previously, but with all systematic parameters set to zero (their central value), the ratio $Q = \frac{L(1)}{L(0)}$ is used as a test statistic. It was found that the sensitivity could be improved by finding values of the heavy flavor *k*-factor and mistag systematic uncertainty parameters that maximized the likelihood. Maximizing other sources of systematic uncertainty was not found to noticeably increase the sensitivity and is computationally expensive.

The cross section is quoted for an assumed value of the top quark mass and can be scaled to give the answer for other values of the mass. Because it is desirable that *p*-value

be independent of any other parameters, the top quark mass is treated as a nuisance parameter in the *p*-value calculation (with a one-sigma uncertainty equal to 2.5 GeV/c^2), so the answer is independent of the value of the top quark mass.

Once this test statistic is chosen, the distribution of $-2 \ln Q$ is plotted. (The logarithm keeps the answers on roughly the same order of magnitude; the factor of –2 is used in deference to convention.) The fraction of events that is less than the measured value is the *p*-value of the measurement.

Because the amount of single top production can fluctuate substantially according to the intrinsic uncertainty of a quantum mechanical process, another question involves the *sensitivity* of the analysis: how well does the analysis perform on average, independent of the fluctuation of the data? The sensitivity can be defined as the *p*-value for which 50% of possible experiments would do better and 50% would do worse: the median expected *p*-value in many pseudo-experiments. This number is obtained by creating many pseudo-experiments with single top production at the Standard Model cross section and plotting the $-2 \ln Q$ distribution of these events. The fraction of events that is less than the median value is the expected sensitivity.

9.5 Systematic uncertainties

Because this analysis relies so heavily on Monte Carlo simulation, a large number of systematic uncertainties must be introduced to make sure that the data and Monte Carlo

are consistent and the answer is trustworthy. Adding a systematic uncertainty increases the expected uncertainty on the cross section but is necessary when there is reason to believe the Monte Carlo is systematically different from the data. There are many sources of possible uncertainty, both rate and shape uncertainties, which are enumerated here.

Cross section

The uncertainty on the theoretical cross section of each background process is treated separately as a source of rate uncertainty. The uncertainties used are listed in Table 13. There is no uncertainty applied for the single-top cross section because this is the quantity being measured.

Scale factors

The uncertainty on the lepton identification scale factor (see page 85), the trigger efficiency (see page 73), and the *b*-tagging scale factor (see page 109) are each included as a source of uncertainty for each sample that uses these factors. The rate uncertainties differ by sample and by channel but are correlated.

Luminosity

Background estimates that rely on theoretical cross sections also have a rate uncertainty on the luminosity calculation (see Table 2). Also included is the uncertainty on the scaling required by the cut on the primary vertex *z*-position (see page 98).

Heavy flavor scale factor

The scaling applied to the heavy flavor sample has a large rate uncertainty that is correlated across all heavy flavor samples (see page 137). This is one of the largest systematic effects in this analysis, though its impact shrinks with increasing integrated luminosity because the likelihood fit is better able to set the background normalization in the data.

Mistag estimate

The total contribution of the mistag is calculated using mistag probabilities described on page 113. These have systematic uncertainties which are also included as a rate uncertainty on the total number of mistagged events.

QCD multijet estimate

The large uncertainty on the QCD multijet fraction, as described on page 134, is included as a rate uncertainty.

Jet energy scale

Improper estimation of the jet energy scale corrections can cause both rates and shapes to change. This uncertainty is estimated by reprocessing all Monte Carlo events with the jet energy scale uncertainties (see page 56 ff.) shifted to their highest value, and again with all uncertainties shifted to their lowest value. The background estimate is recalculated to determine the change in the rates and the resulting EPD templates are used to derive shape

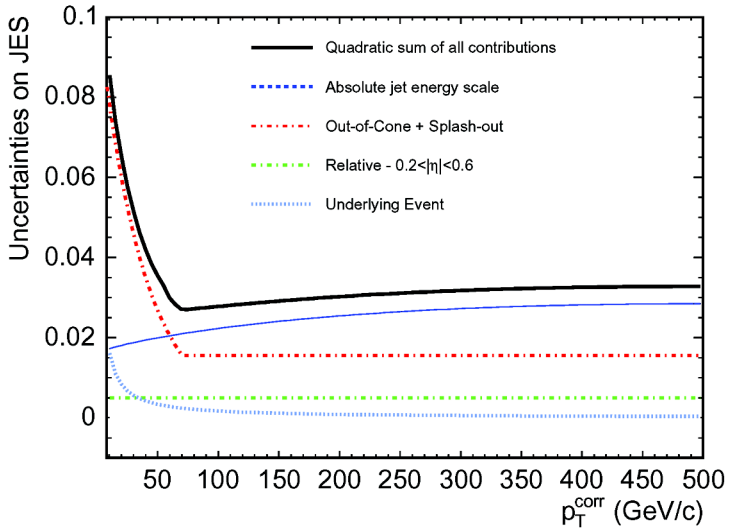

Figure 66. The systematic uncertainty on the jet energy scale as a function of the transverse momentum of the jet. This is the largest single source of systematic uncertainty in this analysis.

uncertainties for all signal and background processes (except for QCD multijet events, which are derived from data). This uncertainty has the largest impact on this analysis of all sources of systematic uncertainty.

Initial state radiation

The PYTHIA showering algorithm deals well with initial state radiation (ISR)—gluons radiated from partons before they interact—but still has some uncertainty. The amount of ISR depends on several physics quantities, including the momentum scale of the interaction and the value of Λ_{QCD}, the energy scale at which perturbative quantum chromodynamics becomes impossible [102]. A systematic uncertainty is applied to both rates and shapes to account for incomplete knowledge of these parameters.

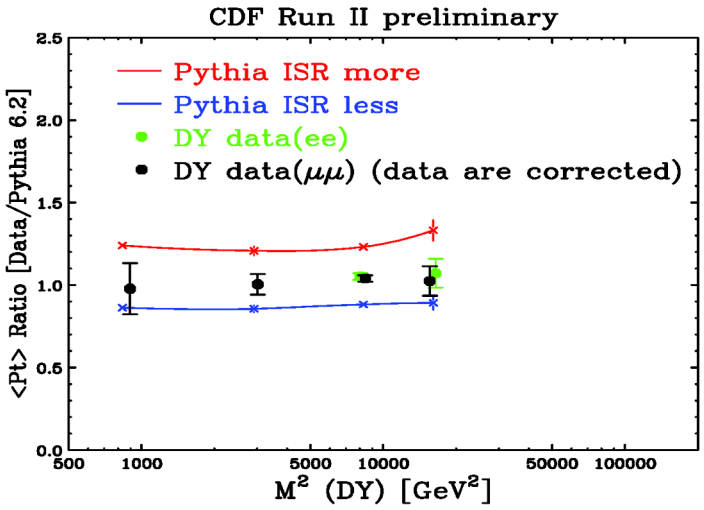

Figure 67. The ratio of transverse momentum between data and Monte Carlo as a function of the invariant mass of the leptons in the event. The systematic uncertainty covers the data points and their statistical uncertainty.

The performance of PYTHIA can be checked by looking at Drell-Yan events: quark pairs that turn into a photon or *Z* boson and then decay into two leptons. This sample has the advantage that it has no final state radiation, which normally would interfere if there were quarks or gluons in the final state. The requirement of two leptons also allows the construction of a very pure sample, and the process' large cross section ensures enough data to make a statistically significantly measurement.

The data match the Monte Carlo events very well, as shown in Figure 67. The systematic uncertainty is constructed by varying three parameters in PYTHIA: the value of Λ_{QCD} is doubled or divided in half; the initial transverse momentum scale is multiplied by four or divided by four; and the hard scattering scale of the shower is multiplied by four.

The increased parameters define one sample, and the decreased parameters define another sample. These two samples are taken to be one standard deviation from the central value and are used to cover the uncertainty of ISR. ISR uncertainties are only applied to signal and the top pair production background. The W + jets background does not need a rate uncertainty because it is scaled to data, and its shape uncertainty is dealt with by the factorization scale uncertainty.

Final state radiation

Unlike ISR, final state radiation (FSR) cannot be measured in isolation at the Tevatron because all interactions begin with quarks, and thus it is impossible to remove ISR contamination. However, the PYTHIA FSR parameters have been tuned at the Large Electron-Positron storage ring (LEP) [103], which has leptons in its initial state and is thus able to measure FSR cleanly. The FSR systematic samples are constructed by varying analogous parameters to the ISR case, except for the hard scattering scale of the shower, which is not applicable to FSR. This creates two new samples which are used to derive rate and shape uncertainties. As with ISR, FSR uncertainties are only applied to signal and the top pair production background.

Parton distribution functions

The parton distribution functions (PDFs) of the protons and antiprotons are not perfectly known. Besides the CTEQ PDFs used in this analysis, there is another set of PDFs

derived by Martin, Roberts, Stirling, and Thorne (MRST) [104]. Leading-order or next-to-leading-order PDFs can be used, various parameters which are derived from experimental data have a finite uncertainty, and the value of the strong coupling constant can be changed. PDF uncertainties are calculated by reweighting events based on different PDF schemes. This reweighting affects both rates and shapes. The changes caused by each of the twenty input variables of the CTEQ PDFs are added in quadrature. In addition, changes are calculated for other versions of CTEQ and MRST and with two different values of the strong coupling constant. If any of these is larger than the uncertainty from the CTEQ parameters, it is also added in quadrature. The final sum gives an event weight which can be used to calculate the change in rate and shape. PDF uncertainties are only calculated for the signal samples.

Renormalization and factorization scale

The ALPGEN event generator used for W + jets events requires the renormalization and factorization scale to be set to solve the divergences caused by gluon splitting. The factorization scale is normally set to the quadrature sum of the W boson mass and the transverse momenta of the jets in the event. The renormalization scale of the strong coupling constant is normally set to the transverse momentum of the interaction vertex.

Because these values are not known, and indeed not physically measurable, since they are artifacts of perturbation theory, an uncertainty is assigned to cover a variety of different possibilities. Both parameters are doubled and halved to create two samples which

are used as a shape uncertainty on the $W + bb$ template. Any effects of this parameter on the rate are ignored because the sample is scaled to the data and this scaling already has a rate uncertainty assigned.

Jet flavor separator

The neural-net jet flavor separator performs very well in its *b*-quark and light-quark control regions once its correction function is applied. The problem lies with the in-between case of charm quarks. It is very difficult to construct a data sample with a high purity of *c*-jets, so it is difficult to test the shape of the *c*-jet distribution. This is accounted for by a systematic uncertainty which envisions "optimistic" and "pessimistic" scenarios.

The optimistic scenario is that the light jets do not need a correction function at all, in which case light jets are even easier to separate from heavy jets. The pessimistic scenario assumes that the *c*-jets are mismodeled as badly as the light jets and thus need the same correction function, which makes both light jets and *c*-jets harder to separate from *b*-jets. The true value almost certainly lies in this range, so the systematic uncertainty covers any effect from the lack of understanding of the data. This systematic uncertainty only affects the shapes of the templates.

Mistag model

Mistagged events are difficult to properly model, so their shape may not be trustworthy. This can be conservatively covered with a systematic shape uncertainty that

uses untagged data weighted by the mistag matrix to construct the mistag template. This uncertainty will absorb any mismodeling in the W + light jets sample by allowing the shape to like untagged data, which is mostly W + light jets.

QCD multijet flavor composition

The QCD multijet flavor composition was estimated in a control region. However, the extrapolation to the signal region may be imperfect, and in addition there is a large statistical uncertainty on measured fraction in the control region. A systematic sample uses the "worst-case" fraction of 60% *b* quarks, 30% *c* quarks, and 10% light quarks. (This is considered the worst case because it makes the QCD multijet sample more signal-like, making it harder to discriminate from the signal.) This only affects the shape of the QCD multijet template.

Mismodeling of jet pseudorapidity

Cross-checks in the untagged region show a mismodeling of the lowest-transverse-energy jet in regions of high positive pseudorapidity. This is of considerable importance because a jet in this region is a key feature of *t*-channel events, and such a mismodeling might exhibit itself as false signal. This effect was studied extensively and was not found to be correlated to any other variable or known effect. It is theorized that there is unmodeled "beam splash" in the far-forward region of the detector which is larger on the east side because the more numerous protons go that direction in elastic scattering. It is also possible

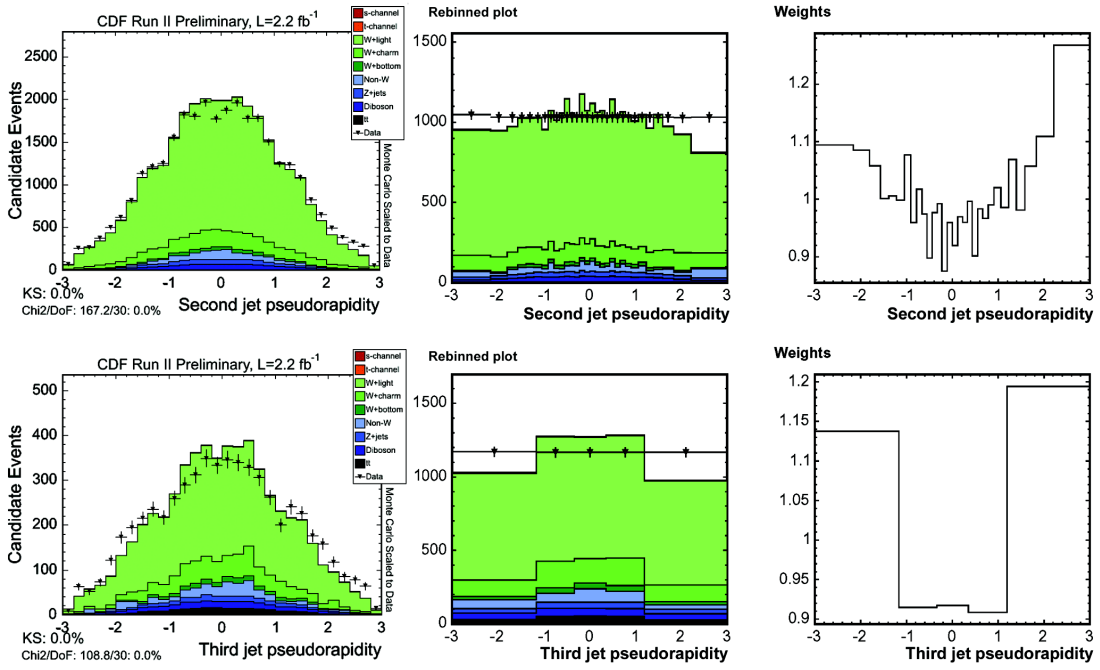

Figure 68. The poorly modeled distributions in the untagged sample of (top) the pseudorapidity of the second jet for two-jet events, and (bottom) the pseudorapidity of the third jet for three-jet events. The histograms (left) are rebinned with variable-width bins (center) to reduce the effects of statistical uncertainty. Dividing the data by the Monte Carlo histograms produces the weight histograms (right), which are then applied to the tagged sample to produce shapes for use as a systematic uncertainty.

that the jet energy scale correction does not fully account for all effects in this region. However, because the effect is not understood, a maximal systematic is assigned: weights are calculated in the untagged sample to force the Monte Carlo to match the data. These weights, applied to the Monte Carlo in the tagged sample, constitute a shape uncertainty.

ΔR mismodeling

Another variable was found similarly mismodeled in the untagged sample: the distance between the two jets in the η–ϕ plane, usually referred to as ΔR. This mismodeling

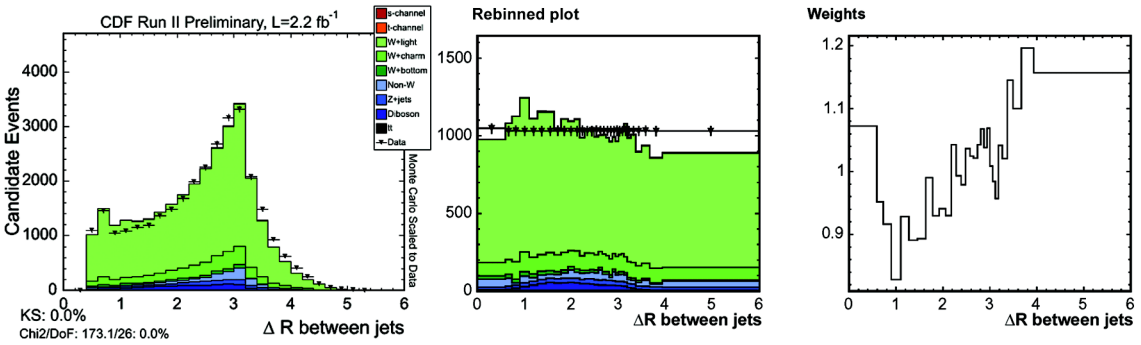

Figure 69. The untagged distribution of the poorly modeled variable ΔR in two-jet events. The histogram (left) is rebinned in variable-width bins (center) to reduce the effect of statistical fluctuations, and a histogram of weights (right) is derived and applied to the tagged sample to create a shape that is used as a systematic uncertainty.

is believed to be caused by a problem in the calculation of gluon splitting in ALPGEN; however, since this is not yet shown convincingly, the difference is used as the basis of a systematic similar to the one constructed for jet pseudorapidity. This is a conservative approach that covers all Monte Carlo and data events.

All systematic uncertainties, once the shape differences have been smoothed, are incorporated into the likelihood as nuisance parameters. All that remains is to look at the EPD distribution of the data and find the value of the cross section which maximizes the likelihood.

Chapter 10

Results and conclusions

Once all matrix elements have been calculated, templates created, cross checks completed, and systematic uncertainties accounted for, the EPD distribution of the data is compared to the Monte Carlo templates (Figure 70). Qualitatively, the data in the most signal-like bins of the most sensitive sample—events with two jets and one tag—fall lower than the expectation, though still statistically consistent with it.

10.1 Cross section measurement

Figure 71 shows the likelihood distribution in the data, with the 68% region marked to show the uncertainty. The measured value is $2.2^{+0.8}_{-0.7}$ pb. This analysis was performed under the assumption that the top mass is 175 GeV/c^2. Other values of the top mass are examined as a cross check. If the top mass is assumed to be 170 GeV/c^2, the measured cross section is 2.0 pb; if the top mass is 180 GeV/c^2, the cross section is 2.3 pb.

10.2 Measurement of $|V_{tb}|$ and limit

The CKM matrix element $|V_{tb}|$ can be calculated very simply. Because the cross section is proportional to the square of V_{tb}, $|V_{tb}|$ is just the square root of the measured cross section divided by the expected cross section from the Standard Model. This gives a result

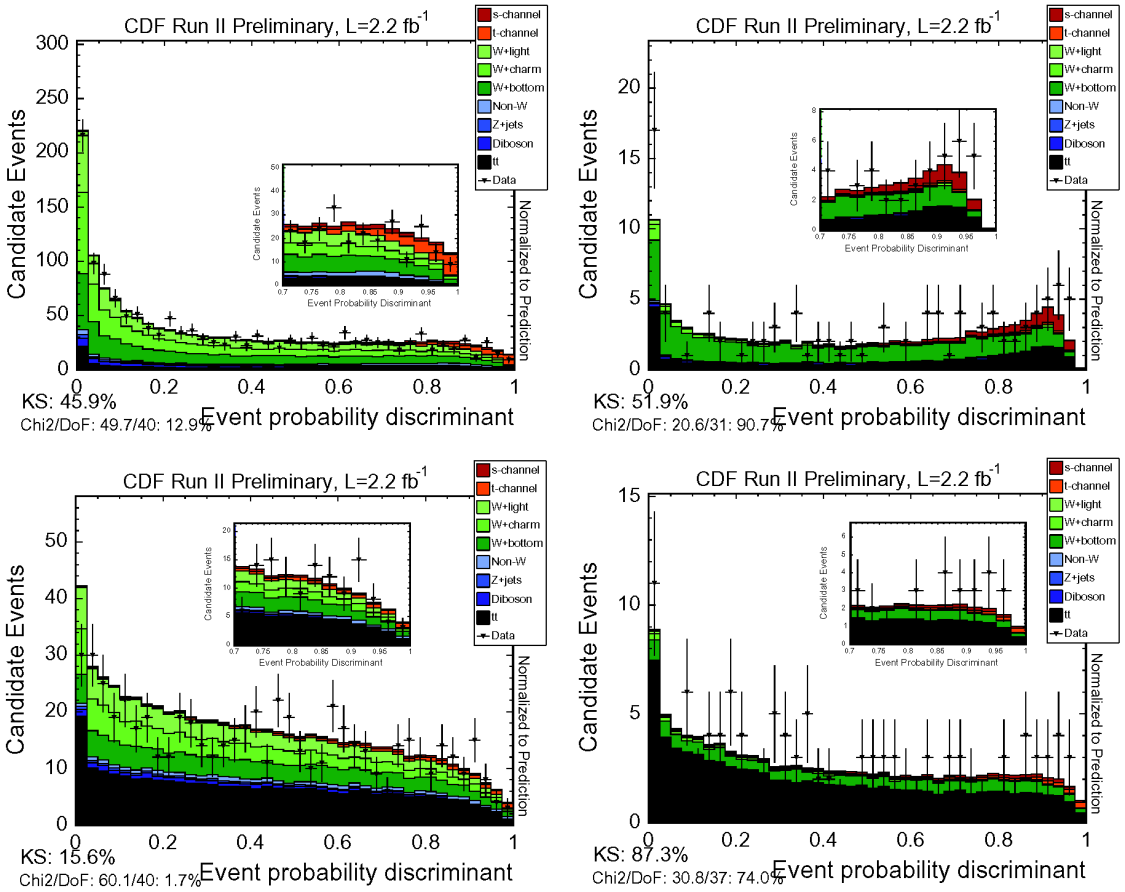

Figure 70. The EPD distribution in Monte Carlo and data in events with (top left) two jets and one *b*-tag, (top right), two jets and two *b*-tags, (bottom left) three jets and one *b*-tag, and (bottom right) three jets and two *b*-tags.

of $|V_{tb}| = 0.88^{+0.16}_{-0.14}(\text{experimental}) \pm 0.07(\text{theoretical})$. Because this measurement relies on the theoretical cross section of electroweak single top production, it must include the uncertainty on the cross section calculation, which is quoted here as the theoretical uncertainty.

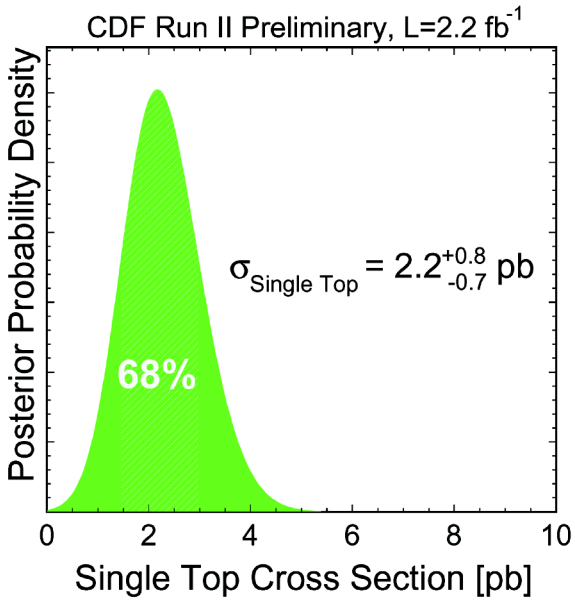

Figure 71. The probability density resulting from the single top cross section measurement.

It is also possible to set a lower limit on $|V_{tb}|$. A flat prior is assumed in $|V_{tb}|^2$, which is proportional to the measured cross section, and it is required to lie between zero and one inclusive. The likelihood must now be modified to include the systematic uncertainties on the single top production cross section and the top mass. This likelihood curve is then integrated from one, the maximum allowed value, until the area covered includes the desired percentage of the total area under the likelihood curve. Conventionally, limits are quoted at 95% confidence, so the curve is integrated until 95% of the area is included. This sets a lower limit on $|V_{tb}|$ of 0.6. Checks with pseudo-experiments show that this limit gives proper frequentist coverage.

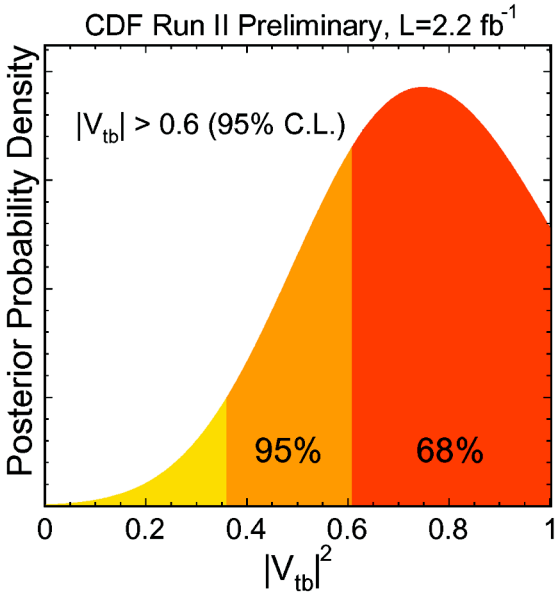

Figure 72. The likelihood in $|V_{tb}|^2$ used to set a 95% lower limit on the value of $|V_{tb}|$.

10.3 Sensitivity and significance

The *p*-value of this measurement is obtained from the distribution of $-2 \ln Q$, plotted in Figure 73. The fraction of events that give a value less than the measured value is the *p*-value of the measurement. This measurement gives a *p*-value of 0.0003. It is customary to interpret this probability as a number of standard deviations from the mean of a Gaussian distribution, integrated from the center of the distribution. In this case, the measured significance corresponds to a 3.4 standard deviations in Gaussian statistics.

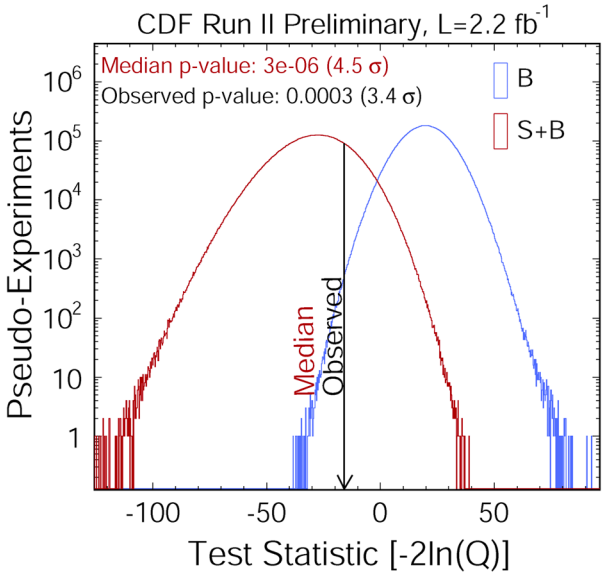

Figure 73. Distribution of the test statistic used to calculate the *p*-value and sensitivity of the single top measurement.

Figure 73 also shows the distribution of events thrown from templates that include the Standard Model contribution of single top production. The sensitivity is the *p*-value of the median of the signal distribution. In this analysis, the sensitivity is 0.000003, which corresponds to 4.5 standard deviations in Gaussian statistics.

10.4 Conclusions

The measurement of the cross section of single top quark production at CDF is a challenging task, requiring a good understanding of the detector, careful simulation, detailed studies of background processes, sophisticated techniques for separating signal from background, and a rigorous statistical treatment. The result of this analysis is one of the world's most sensitive measurements of the cross section and of $|V_{tb}|$. This analysis

measures a cross section of $2.2^{+0.8}_{-0.7}$ pb, which corresponds to a value of $|V_{tb}| = 0.88^{+0.16}_{-0.14}(\text{experimental}) \pm 0.07(\text{theoretical})$. The probability that this result comes from a fluctuation from the background-only hypothesis (*p*-value) is 0.0003, which corresponds to 3.4 standard deviations in Gaussian statistics. The expected (median) *p*-value calculated in pseudo-experiments for this analysis is 0.000003, which corresponds to 4.5 standard deviations in Gaussian statistics.

The single top cross section has now been measured to good precision, and these results have been corroborated by other measurements at CDF [105] and at D0 [106]. The precision will continue to improve with more data: at the time of this writing, the Tevatron has delivered more than 4 fb^{-1} of integrated luminosity, of which nearly 3 fb^{-1} are available for analysis by the CDF and D0 experiments. The amount of data collected is such that single top quark production measurements are beginning to be limited by systematic uncertainties instead of statistical uncertainties.

Future analyses will focus on reducing systematic uncertainties, some of which are overly conservative. In addition, the sample is becoming large enough to measure *s*-channel and *t*-channel cross sections separately, search for evidence of associated *W* boson production, and measure the polarization of the top quarks produced in this channel (predicted by the Standard Model to be nearly 100%). Many useful physics measurements related to electroweak single top quark production remain!

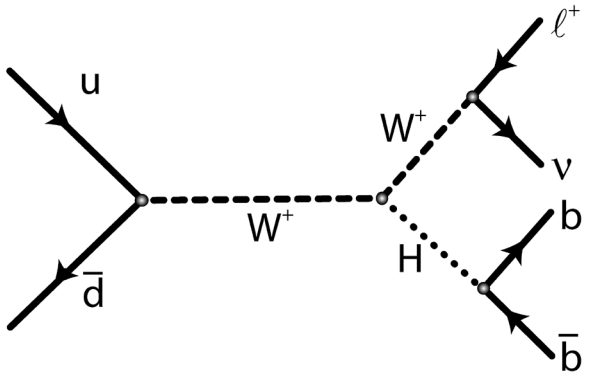

Figure 74. The Feynman diagram for *WH* production. It has the same final state as single top production.

10.5 Outlook: Searching for the Higgs boson

The matrix element method used in this analysis can be employed in other measurements that have a large background and a small signal contribution to the sample. One very useful application of this method is in the search for the Higgs boson.

The Higgs boson is a scalar field predicted by Peter Higgs [107] that would account for the masses of the *W* and *Z* bosons, as well as other particles, through electroweak symmetry breaking. Although it has a strong theoretical grounding, it has not yet been observed because of its small coupling. A low-mass Higgs boson (with a mass of 100–130 GeV) is most easily seen at the Tevatron in the *WH* production mode. This process has the same final state as single top, but its predicted cross section is ten to a hundred times smaller than the *s*-channel cross section.

Because the final state of *WH* production is the same as that of single top production, it is simple to use the same event selection, background estimate, and analysis method as the

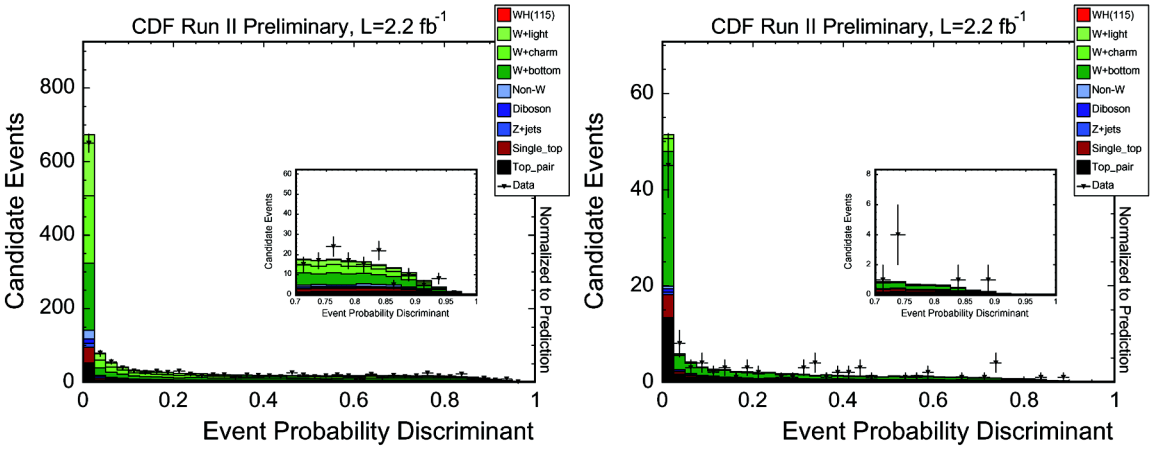

Figure 75. Examples of the EPDs constructed for the *WH* search. These distributions were optimized for a Higgs mass of 115 GeV/c^2. (left) The distribution for events with exactly one *b*-tag, and (right) the distribution for events with more than one *b*-tag. The *WH* signal, which has a Standard-Model cross section of about 0.15 pb, is not visible.

single top production analysis to search for the Higgs boson. This analysis is conducted on the same data set, adding a matrix element for *WH* production calculated at different mass points (100, 115, 120, 130, 140, and 150 GeV/c^2). The EPD now has the *WH* matrix element in the numerator and single top production as one of the backgrounds. A different EPD is constructed for each Higgs mass, and each EPD has its parameters tuned to maximize sensitivity.

There is not enough data currently to see a discernable signal, but upper limits can be set on the Higgs boson cross section. This limit is set in the same way as the $|V_{tb}|$ lower limit, except that the integration begins from zero. Expected limits are calculted by finding the limit in many pseudo-experiments and taking the median. The final limits are shown in Figure 76. The limit at 115 GeV is eight times the Standard Model prediction, which is

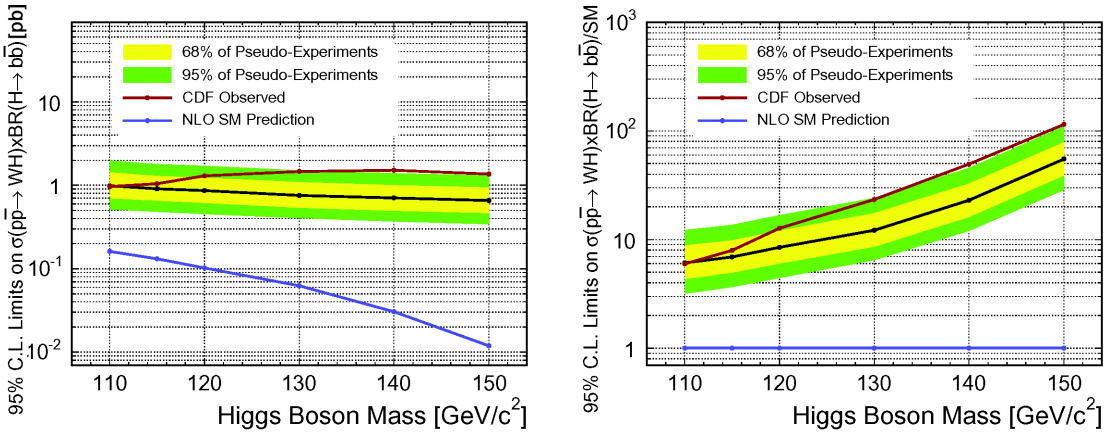

Figure 76. 95% upper limits set on the *WH* production cross section, (left) in terms of cross section and (right) as a ratio to the Standard Model prediction.

currently one of the best limits on a low-mass Higgs boson cross section set by a single analysis. The methodology developed for this single top search has many applications in other branches of particle physics.

Bibliography

[1] For example, "Termites eat through wood two times faster when listening to rock music," http://www.snapple.com/realfacts/.

[2] G. Altarelli, "Status of the Standard Model and Beyond," arvix.org:hep-ph/030605.

[3] J. Alwall, *et al.*, "Is $V_{tb} \simeq 1$?," arxiv.org:hep-ph/0607115v1.

[4] For an overview of the Standard Model, see, for example, D. Griffiths, *Introduction to Elementary Particles*, John Wiley & Sons (1987) pp. 46–48, 55–56.

[5] For a simple overview of local gauge invariance, the Standard Model, and the Higgs mechanism, see *ibid*., pp.344–358.

[6] *Ibid*., pp. 55–56.

[7] W.-M. Yao et al. [Particle Data Group], J. Phys. G **33**, 1 (2006) and 2007 partial update for edition 2008, http://pdg.lbl.gov.

[8] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*, Westview Press (1995) pp. 479–480.

[9] *Ibid.*, pp. 723–724.

[10] See, for example, A. Zee, *Quantum Field Theory in a Nutshell*, Princeton University Press (2003) pp. 41–60.

[11] An excellent qualitative discussion of renormalization can be found in *ibid.*, pp. 145–151.

[12] *Op. cit.* Peskin, pp. 553–554.

[13] For a derivation see *ibid*., pp. 99–108.

[14] For example, *Op. cit*. Griffiths, pp. 198–199.

[15] *Loc. cit.* Yao.

[16] F. Abe *et al.* [CDF Collaboration], "Observation of Top Quark Production in p anti-p Collisions with the Collider Detector at Fermilab," Phys. Rev. Lett. **74** (1995) 2626;

S. Abachi *et al*. [D0 Collaboration], "Observation of the Top Quark," Phys. Rev. Lett. **74** (1995) 2632.

[17] V. Abazov *et al.* [D0 Collaboration], "Simultaneous Measurement of the Ration B($t \to Wb$) / B($t \to Wq$) and the Top Quark Pair Production Cross Section with the D0 Detector at $\sqrt{s} = 1.96$ TeV," accepted in Phys. Rev. Lett., arXiv.org:hep-ex/0801.1326.

[18] G. Mahlon, "Improved Spin Basis for Angular Correlation Studies in Single Top Quark Production at the Tevatron," Phys. Rev. D **55** (1997) 7249.

[19] M. Carena *et al.*, "Report of the Higgs Working Group of the Tevatron Run 2 SUSY/Higgs Workshop," arXiv.org:hep-ph/0010338v2.

[20] For example, S. Weinberg, Phys. Rev. Lett. **43** (1979) 1566.

[21] Z. Sullivan, "Understanding Single-Top-Quark Production and Jets at Hadron Colliders," Phys. Rev. D **70** (2004) 114012.

[22] D. Perkins, *Introduction to High Energy Physics*, Cambridge University Press (2000) p. 343.

[23] For example, see A.Streun, "Practical Guidelines for Lattice Design," Swiss Light Source internal report SLS-TME-TA-1999-0014, http://slsbd.psi.ch/pub/cas/cas/node38.html.

[24] http://lhc.web.cern.ch/lhc/.

[25] http://www.cern.ch/.

[26] *Op. cit.* Perkins, p.342.

[27] R. Wilson, "The Tevatron," Phys. Today **30**, 10 (1977) 23.

[28] J. Cockroft, E. Walton, "Experiments with High Velocity Positive Ions. II. The Disintegration of Elements by High Velocity Protons," Proc. of the Royal Soc. of London A **137** (1932) 229.

[29] L. Alvarez, Phys. Rev. **70** (1942) 799.

[30] http://www-bd.fnal.gov/ntf/ntf_home.html.

[31] http://www-boone.fnal.gov/.

[32] http://www-numi.fnal.gov/

[33] J. Marriner, "Stochastic Cooling Overview," Nucl. Instrum. Meth. A **532** (2004) 11.

[34] G. Parzen, "Theory of Electron Cooling using Electron Cooling as an Intrabeam Scattering Process," arxiv.org:physics/0609076.

[35] For a complete but outdated overview of the detector, see the *CDF II Technical Design Report*, available at http://www-cdf.fnal.gov/upgrades/tdr/tdr.html, along with F. Abe, *et al.* [CDF Collaboration], "The CDF Detector: An Overview," Nucl. Instrum. Meth. A **271** (1988) 387.

[36] Principles of silicon detectors are discussed, for example, in K. Kleinknecht, *Detectors for Particle Radiation*, University Press (1998) p. 62.

[37] J. Antos, *et al.*, "The SVX II Silicon Vertex Detector Upgrade at CDF," Nucl. Instrum. Meth. A **383** (1996) 13.

[38] T. Affolder, *et al.*, "CDF Central Outer Tracker," Nucl. Instrum. Meth. A **526** (2004) 249.

[39] For a discussion of the principles of drift chambers, see *op. cit.* Kleinknecht, pp. 78–81.

[40] D. Acosta, *et al.*, "A Time-of-Flight Detector in CDF-II," Nucl. Instrum. Meth. A **518** (2004) 605.

[41] H. Minemura, *et al.*, "Construction and Testing of a 3 m Diameter × 5 m Superconducting Solenoid for the Fermilab Collider Detector Facility (CDF)," Nucl. Instrum. Meth. A **238** (1985) 18.

[42] *Op. cit.* Perkins, p. 372.

[43] L. Balka, *et al.*, "The CDF Central Electromagnetic Calorimeter," Nucl. Instrum. Meth. A **267** (1988) 280; S. Bertolucci, *et al.*, "The CDF Central and Endwall Hadron Calorimeter," Nucl. Instrum. Meth. A **267** (1988) 301.

[44] R. Oishi, *et al.*, "New CDF End-Plug Calorimeter," Nucl. Instrum. Meth. A **453** (2000) 227.

[45] G. Ascoli, *et al.*, "The CDF Central Muon Detector," Nucl. Instrum. Meth. A **268** (1988) 41; T. Dorigo, "The Muon System Upgrade for the CDF II Experiment," Nucl. Instrum. Meth. A **461** (2001) 560.

[46] D. Acosta, *et al.*, "The CDF Cherenkov Luminosity Monitor," Nucl. Instrum. Meth. A **461** (2001) 540.

[47] D. Amidei, *et al.*, "The CDF Trigger," Nucl. Instrum. Meth. A **265** (1988) 326.

[48] A. Abulencia, *et al.*, "The CDF II Extremely Fast Tracker Upgrade," Nucl. Instum. Meth. A **581** (2007) 482.

[49] J. Adelman, *et al.*, "The Silicon Vertex Trigger Upgrade at CDF," Nucl. Instrum. Meth. A **572** (2007) 361.

[50] http://www.visitmonaco.com/index.cfm?fuseaction=Page.viewPage&pageId=734.

[51] J. Pulplin, *et al.*, "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis," Jour. High En. Phys. **0207** (2002) 012.

[52] For example, the MC@NLO generator: S. Frixione and B. Webber, "Matching NLO QCD Computations and Parton Shower Simulations," Jour. High En. Phys. **0206** (2002) 029.

[53] See, for example, V. Barger and R. Phillips, *Collider Physics*, Addison-Wesley (1987).

[54] T. Sjöstrand, *et al.*, "High-Energy-Physics Event Generation with PYTHIA 6.1," Comput. Phys. Commun **135**, (2001) 238.

[55] J. Alwall, *et al*., "MadGraph/MadEvent v4: The New Web Generation," J. High En. Phys. **0709** (2007) 028.

[56] M. Mangano, *et al*., "ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions," J. High En. Phys. **0301** (2003) 001.

[57] A detailed overview of PYTHIA's showering routines can be found in the PYTHIA manual at http://home.thep.lu.se/~torbjorn/pythia/lutp0613man2.pdf.

[58] http://www.lns.cornell.edu/public/CLEO/soft/QQ/qq_expert.html.

[59] See the CLEO collaboration web page at http://www.lepp.cornell.edu/Research/EPP/CLEO/.

[60] S. Jadach, J.H. Kühn and Z. Wąs, "TAUOLA - A Library of Monte Carlo Programs to Simulate Decays of Polarized τ Leptons." Comput. Phys. Commun. **64** (1991) 275.

[61] S. Agostinelli *et al.*, "GEANT 4—A Simulation Toolkit," Nucl. Instrum. Meth. A **506** (2003) 250.

[62] G. Grindhammer and S. Peters, "The Parameterized Simulation of Electromagnetic Showers in Homogeneous and Sampling Calorimeters," arxiv.org:hep-ex/001020.

[63] For more information on the jet energy scale corrections at CDF, see A. Bhatti, *et al*., "Determination of the Jet Energy Scale at the Collider Detector at Fermilab," Nucl. Instrum. Meth. A **566** (2006) 375.

[64] F. Abe, *et al.* [CDF Collaboration], "Measurement of the Antiproton-Proton Total Cross Section at $\sqrt{s}=$ 546 and 1800 GeV," Phys. Rev. D **50** (1994) 5550.

[65] For example, *op. cit.* Perkins, p. 349.

[66] *Loc. cit.* Alwall.

[67] *Loc. cit.* Sullivan.

[68] *Loc. cit.* Sullivan.

[69] More details on this work can be found in J. Lück, Diploma thesis, University of Karlsruhe (2006).

[70] Z. Sullivan, "Understanding Single-Top-Quark Production and Jets at Hadron Colliders," arXiv.org:hep-ph/0408049v2; M. Cacciari, *et al.*, "The $t\bar{t}$ Cross-section at 1.8 and 1.96 TeV: A Study of the Systematics Due to Parton Densities and Scale Dependence," Journ. High En. Phys. **0404** (2004) 068.

[71] J. Campbell, R. Ellis, "Update on Vector Boson Pair Production at Hadron Colliders," Phys. Rev. D **60** (1999) 113006.

[72] T. Aaltonen, *et al.* [CDF Collaboration], "Measurement of Inclusive Jet Cross Sections in $Z/\gamma^*(\rightarrow e^+e^-)$+jets Production in $p\bar{p}$ Collisions at $\sqrt{s}=1.96$ TeV," Phys. Rev. Lett. **100** (2008) 102001.

[73] B. W. Harris, *et al*. "Fully Differential Single-Top-Quark Cross Section in Next-To-Leading Order QCD." Phys. Rev. D **66** (2002) 054024.

[74] *Loc. cit.* Sullivan.

[75] Loc. cit. Sullivan.

[76] T. Aaltonen, *et al.* [CDF Collaboration], "First Measurement of the Production of a W Boson in Association with a Single Charm Quark in $p\bar{p}$ Collisions at $\sqrt{s}=1.96$ TeV," Phys. Rev. Let. **100** (2007) 091803.

[77] D. Acosta, *et al.* [CDF Collaboration], "Search for Electroweak Single-Top-Quark Production in $p\bar{p}$ Collisions at $\sqrt{s}=1.96$ TeV," Phys. Rev. D **71** (2005) 012005.

[78] C. P. Yuan, "A New Method to Detect a Heavy Top Quark at the Tevatron," Phys. Rev. D **41** (1990) 42.

[79] F. Canelli, *Helicity of the W Boson in Single-Lepton $t\bar{t}$ Events*, Ph.D. thesis, University of Rochester (2003).

[80] For example, *Op. cit.* Griffiths, pp. 198–199.

[81] *Loc. cit.* Sullivan.

[82] *Loc. cit.* Alwall.

[83] H. Murayama, I. Watanabe, and K. Hagiwara, "HELAS: Helicity Amplitude Subroutines for Feynman Diagram Evaluations," KEK Report 91-11 (1992). The documentation can be found, for example, at home.fnal.gov/~mrenna/helas.ps.

[84] A. van Doren and L. de Ridder, "An Adaptive Algorithm for Numerical Integration Over an *n*-Dimensional Cube," J. Comput. Appl. Math. **2** (1976) 207.

[85] The CERN Program Library is a package of libraries and modules for use in particle physics analyses. More information and documentation can be found at http://cernlib.web.cern.ch/cernlib/.

[86] A. Genz and A. Malik, "Remarks on Algorithm 006: An Adaptive Algorithm for Numerical Integration over an *N*-Dimensional Rectangular Region," J. Comput. Appl. Math. **6** (1980) 295; implemented as CERNLIB algorithm D120, documented at http://wwwasdoc.web.cern.ch/wwwasdoc/shortwrupsdir/d120/top.html.

[87] R. Brun and F. Rademakers, "ROOT—An Object Oriented Data Analysis Framework," Nucl. Instrum. Meth. A **389** (1997) 81. See also http://root.cern.ch/.

[88] J. Friedman and M. Wright, "A Nested Partitioning Procedure for Numerical Multiple Integration," ACM Trans. Math. Software **7** (1981) 76; implemented as CERNLIB algorithm D151, documented at http://wwwasdoc.web.cern.ch/wwwasdoc/shortwrupsdir/d151/top.html.

[89] T. Hahn, "CUBA—A Library for Multidimensional Numerical Integration," Comput. Phys. Commun. **168** (2005) 78.

[90] J. Koksma, "Een Algemeene Stelling Uit de Theorie der Gelijkmatige Verdeeling Modulo 1," Mathematica B **11** (1943) 7; E. Hlawka, "Funktionen von Beschränkter Variation in der Theorie der Gleichverteilung," Ann. Mat. Pura Appl., **54** (1961) 325; see discussion, for example, in F. Hickernell, "Koksma–Hlawka Inequality," *Encyclopedia of Statistical Sciences*, John Wiley & Sons (2006).

[91] For more on the jet flavor separator, see S. Richter, *Search for Electroweak Single Top-Quark Production with the CDF II Experiment*, Ph.D. thesis, University of Karlsruhe (2007).

[92] M. Feindt and U. Kerzel, "The NeuroBayes neural network package," Nucl. Instrum. Meth. A **559** (2006) 190.

[93] T. Kwok and D. Yeung, "Bayesian Regularization in Constructive Neural Networks," in Proceedings of the International Conference on Artificial Neural Networks (1996) 557.

[94] R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*, MIT Press (1998) section 8.2.

[95] *Loc. cit.* Canelli.

[96] J. Neyman and E. Pearson, "On the Problem of the Most Efficient Tests of Statistical Hypotheses," Phil. Trans. of the Royal Soc. Of London A **231** (1933) 289.

[97] A discussion of Bayesian statistics can be found, for example, in B. Efron, "Why Isn't Everyone a Bayesian?" Am. Stat. **40** (1986) 11, and R. Cousins, "Why Isn't Every Physicist a Bayesian?" Am. J. Phys. **63** (1995) 398.

[98] R. Cousins, "Treatment of Nuisance Parameters in High Energy Physics, and Possible Justifications and Improvements in the Statistics Literature," PhyStat05: Statistical Problems in Particle Physics, Astrophysics and Cosmology (2005).

[99] A discussion of frequentist statistics can be found in *loc. cit.* Efron.

[100] W. Eadie, *et al.*, *Statistical Methods in Experimental Physics*, North-Holland (1971), p. 278.

[101] *Loc. cit.* Neyman.

[102] *Loc. cit.* Sjöstrand and PYTHIA manual.

[103] An overview of the LEP experiment can be found at http://sl-div.web.cern.ch/sl-div/history/lep_doc.html.

[104] A. Martin, *et al*., "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis," Journ. High En. Phys. **0207** (2002) 012.

[105] CDF Collaboration, "Multivariate Likelihood Search for Single-Top-Quark Production with 2.2 fb^{-1}," CDF public note 9221, http://www-cdf.fnal.gov (2008); CDF Collaboration, "Search for Electroweak Single Top-Quark Production using Neural Networks with 2.2 fb^{-1} of CDF II data," CDF public note 9217, http://www-cdf.fnal.gov (2008); CDF Collaboration, "Combination of CDF Single Top Quark Searches with 2.2 fb^{-1} of Data," CDF public note 9251, http://www-cdf.fnal.gov (2008); CDF Collaboration, "Search for Single Top Quark Production using Boosted Decision Trees in 2.2 fb^{-1} of CDF Data," CDF public note 9313, http://www-cdf.fnal.gov (2008); paper currently in preparation for submission to Phys. Rev. Lett.

[106] V. Abazov, *et al.* [D0 Collaboration], "Evidence for Production of Single Top Quarks and First Direct Measurement of $|V_{tb}|$," Phys. Rev. Lett. **98**, (2007) 181802; V. Abazov, *et al*. [D0 Collaboration], "Evidence for Production of Single Top Quarks," accepted in Phys. Rev. D, arXiv.org:hep-ex/0803.0739.

[107] P. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," Phys. Rev. Lett. **13** (1964) 508.